MEASUREMENT SCIENCE REVIEW, **17**, (2017), No. 3, 100-107

DE GRUYTER OPEN

MEASUREMENT SCIENCE REVIEW

Journal homepage: http://www.degruyter.com/view/j/msr

ISSN 1335-8871

New Discrete Fibonacci Charge Pump Design, Evaluation and Measurement

David Matoušek[1], Jiří Hospodka[1], Ondřej Šubrt[2, 1]

[1]*Department of Circuit Theory, FEE CTU in Prague, Technicka 2, 16627, Prague, Czech Republic, matoudav@fel.cvut.cz*
[2]*ASICentrum, a company of the Swatch Group, Novodvorska 994, 14221, Prague, Czech Republic*

This paper focuses on the practical aspects of the realisation of Dickson and Fibonacci charge pumps. Standard Dickson charge pump circuit solution and new Fibonacci charge pump implementation are compared. Both charge pumps were designed and then evaluated by LTspice XVII simulations and realised in a discrete form on printed circuit board (PCB). Finally, the key parameters as the output voltage, efficiency, rise time, variable power supply and clock frequency effects were measured.

Keywords: Dickson charge pump, Fibonacci charge pump, voltage gain, output series resistance, rise time, efficiency.

1. INTRODUCTION

Charge pumps are DC/DC converters that produce a voltage higher than supply voltage or a negative voltage. Charge pumps are suitable for a lower value of the output current and take advantage of having no inductive storage elements, whereas, conventional DC/DC converters based on inductors or transformers are more suitable for a higher power. A standard variant of a charge pump is the Dickson charge pump. The Dickson charge pump is efficient, but it produces a relatively small voltage gain. Thus, the Dickson charge pump is useful for lower output to input voltage ratios. A Fibonacci charge pump is a charge pump variant with the voltage gain that is gradually increased over pump stages. On the other hand, the Fibonacci charge pump circuit is more complex than the Dickson charge pump circuit solution.

The Dickson Charge Pump (DCP) is a well-known variant of a charge pump [1]. The schematic diagram is shown in Fig.1. The design equations for DCP are summarised in [2].

The differential voltage ΔV between nodes n and $n+1$ is

$$\Delta V = V_{n+1} - V_n = V_S - V_D, \tag{1}$$

where V_S is the voltage swing at each node due to capacitive coupling from the clock [2], V_D is the diode forward voltage.

The optimal value of the voltage swing equals to the amplitude of clock. But the stray capacitance of a node reduces voltage swing [1] as follows

$$V_S = \left(\frac{C_T}{C_T + C_S}\right) \cdot V_{CLK}, \tag{2}$$

where V_S is the voltage swing, C_T is the transfer capacitance, C_S is the stray capacitance, V_{CLK} is the amplitude of clock.

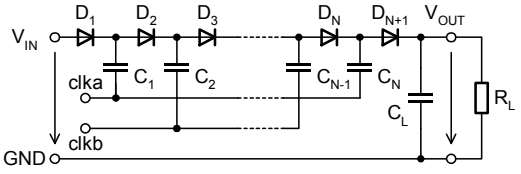

Fig.1. Schematic diagram of the Dickson Charge Pump.

The no-load output voltage applies here according to [1]

$$V_O = V_{IN} + N \cdot (V_S - V_D) - V_D, \tag{3}$$

where V_O is the no-load output voltage, V_{IN} is the input voltage, N is the number of stages, V_S is the voltage swing, V_D is the diode forward voltage drop.

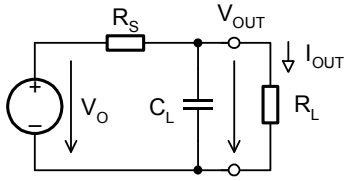

Fig.2. The equivalent circuit of DCP.

The equation (3) assumes the no-load output. The effect of the load current is described by [1] (see Fig.2.).

DOI: 10.1515/msr-2017-0013

$$V_{OUT} = V_O - I_{OUT} \cdot R_S , \tag{4}$$

where V_{OUT} is the output voltage at load, V_O is the no-load output voltage, I_{OUT} is the load current ($I_{OUT} \geq 0$), R_S is the output series resistance of the charge pump.

Reference [1] defines the output resistance of the charge pump as the dependency on the number of stages N, transfer capacitance C_T, stray capacitance C_S and clock frequency f

$$R_S = \frac{N}{(C_T + C_S) \cdot f} \cdot \tag{5}$$

The validity of (5) is limited by the finite resistance of the diodes in use and also the finite output resistance of clock drivers for generating clka and clkb signals [3]. Equation (5) assumes that influence of the resistance of diodes and clock drivers is sufficiently small in the relationship with the equivalent resistance of the transfer capacitors. This condition is usually granted for a low capacitance of the transfer capacitors. But for relatively high values of the transfer capacitance, (5) becomes invalid.

2. SUBJECT & METHODS

The design of the Dickson charge pump and the Fibonacci charge pump was carried out and evaluated by simulations.

A. DCP design and evaluation

The design rules for DCP are summarised in [2] and illustrated by Fig.3. We assume these charge pump specifications:

- power supply voltage $V_{IN} = 3$ V,
- the minimal steady-state output voltage $V_{OUT} = 30$ V at the output current $I_{OUT} = 1$ mA,
- the maximal ripple voltage of the output $V_R = 15$ mV,
- the maximal rise time of the output $t_R = 65$ ms.

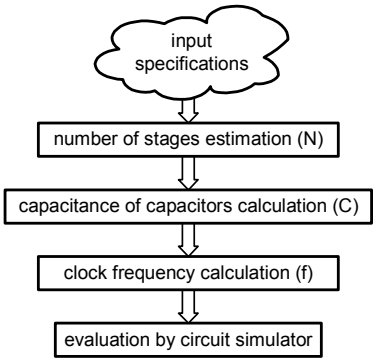

Fig.3. Design-flow diagram.

At first, the number of stages of the Dickson charge pump was estimated by (3) and (4). Ideally, we assume $V_S = V_{IN}$, $V_D = 0$, $R_S = 0$ then the required number of stages is (6). Thus, the number of stages estimation is $N = 9$. As it is evident, the number of stages must be increased. We set $N = 11$.

$$N = \frac{V_{OUT}}{V_{IN}} - 1 , \tag{6}$$

Secondly, the capacitance of capacitors was calculated from known values of the output voltage and current and required rise time of the output (7) [2]. The calculated value is used for the transfer and load capacitors ($C = C_T = C_L$). We calculate $C = 2.17$ μF and then we set $C = 2.2$ μF.

$$C = \frac{I_{OUT} \cdot t_R}{V_{OUT}} , \tag{7}$$

where C is the load and transfer capacitance, V_{OUT}, I_{OUT}, t_R are the voltage, current, and rise time of the output.

As the third step, the clock frequency was calculated from known values of the output current, output ripple voltage, and load capacitance (8) [1]. We calculate $f = 30.3$ kHz and then we set $f = 33$ kHz.

$$f = \frac{I_{OUT}}{V_R \cdot C_L} , \tag{8}$$

where f is the clock frequency, I_{OUT} is the output current, V_R is the ripple voltage of the output, C_L is the load capacitance.

Finally, we select the transistors and diodes. All design parameters are listed in Table 1.

Table 1. Input design parameters and result device parameters.

Parameter	**Value or device**
Output voltage	$V_{OUTmin} = 30$ V ($I_{OUT} = 1$ mA).
Rise time of output	$t_{Rmax} = 65$ ms.
Ripple voltage	$V_{Rmax} = 15$ mV.
Clock frequency	$f = 33$ kHz.
Supply voltage	$V_{IN} = 3$ V.
Capacitances	$C_L = C_T = 2.2$ μF.
NMOS transistor	2N7002 ($V_{DSS} = 60$ V, $V_{GS(th)} = 2.1$ V).
PMOS transistor	BSS84 ($V_{DSS} = -50$ V, $V_{GS(th)} = -1.7$ V).
Schottky diode	PMEG4010BEA ($V_{RRM} = 40$ V, $V_D = 155$ mV).

The final DCP schematic diagram is shown in Fig.4. This circuit solution uses one common clock signal CLK only. The clock signals clka and clkb are derived from the CLK by two inverters M_1, M_2 and M_3, M_4. Thus, these clock signals are overlapped. This solution is easier than a generation of non-overlapped clock signals. At this point, overlapping is not a key factor for the DCP function (this phenomenon does not kill the DCP voltage gain).

The proposed DCP was simulated in LTspice XVII from Linear Technology Corporation. Results from simulations are the output no-load voltage $V_O = 34.90$ V and output voltage $V_{OUT} = 32.78$ V at load ($I_{OUT} = 1$ mA). The rise time of the output is $t_R = 37.07$ ms, and the ripple voltage of the output is $V_R = 7.23$ mV (p-p).

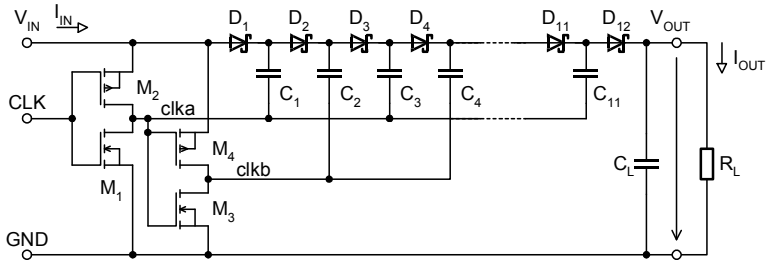

Fig.4. Schematic diagram of the 11-stage DCP.

B. Fibonacci Charge Pump principles

Fibonacci Charge Pump (FCP) [4], [5] is a voltage multiplier with a gradually increasing voltage gain of the stages. The voltage gain of the stage is defined as a Fibonacci number (the Fibonacci sequence is: 1, 1, 2, 3, 5, 8, 13, 21, …).

Schematic diagram of the Fibonacci charge pump is shown in Fig.5.

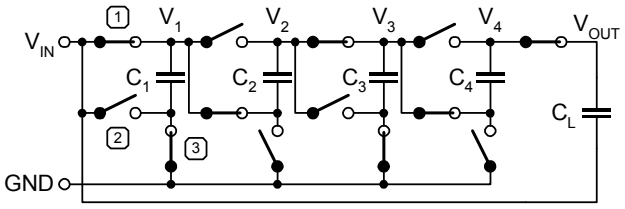

Fig.5. Schematic diagram of the 4-stage Fibonacci charge pump.

The voltages of the individual nodes V_1 to V_4 in periodic steady-state are gradually shifted about multiple of the voltage gain of the first stage. Thus, the 4-stage FCP produces the no-load output voltage $V_O = V_{IN} + 7 \cdot V_{IN}$. Generally, the no-load output voltage V_O is

$$V_O = V_{IN} + V_{IN} \cdot \sum_{n=1}^{N} F_n , \tag{9}$$

where N is the number of stages, F_n is the Fibonacci number of the nth order ($F_1 = 1$, $F_2 = 1$, for $n \geq 3$: $F_n = F_{n-1} + F_{n-2}$).

The effect of the load current is similar as in DCP and it is described in Fig.2. and by the equation (4). The output series resistance of FCP is discussed in [6].

C. Proposed FCP realisation

For a given case according to Table 1., the number of FCP stages must be set to N = 5, because the ideal no-load output voltage for the 5-stage FCP is $V_O = 39$ V. A lower value of the number of stages is not sufficient (e.g. for N = 4 the no-load output voltage is $V_O = 24$ V only) because the required output voltage at the load is 30 V minimally. The values and types of devices according to Table 1. are unchanged for FCP realisation. Thus, we can compare key parameters of DCP versus FCP.

Realisation of the Fibonacci charge pump is more complicated than the Dickson charge pump circuit solution.

The key problem is that FCP uses two floating switches for each stage. For the first stage from Fig.5. these switches are marked as 1 and 2. The switch #1 can be realised as a diode, but the switch #2 must be realised as a transistor. This high-side switch must be realised as a PMOS transistor. This solution is more suitable than driving an NMOS high-side switch. The switch #3 can be realised as an NMOS transistor that works as a low-side switch. To summarise, the switches for the first stage are realised by diode D_1 (switch #1), transistor M_{1b} (switch #2) and transistor M_{1a} (switch #3), see Fig.6.

The second problem of FCP realisation is generating of a driving signal for the switches #2 and #3 for the next stage. A driver for the next stage must be supplied from the output of a current stage and inverts a clock signal to the next stage. The optimal solution of this problem is an auxiliary inverter [7] that is supplied from the output of a current stage and driven from the clock signal of a current stage. This inverter generates inverted and voltage shifted clock signal for the next stage. For example, the second stage of FCP is driven by inverter M_{1c}, M_{1d}, see Fig.6.

The presence of the auxiliary inverter implies the fact that a shoot-through current arises. The intermediate nodes are discharged by this shoot-through current. Thus, the power consumption is increased. This problem may be solved by a more complex architecture of the inverter with an auxiliary current limiter.

The second problem of the implemented auxiliary inverter is the propagation delay. The propagation delay of inverters is gradually increased from the input to the output of the charge pump. The timing discrepancy between the stages may cause a loss of a charge. Thus, the clock period should be set sufficiently long related to this propagation delay.

Proposed new circuit solution of FCP according to Fig.6. was verified by simulation in LTspice XVII. Results from simulations are the output no-load voltage $V_O = 35.00$ V, output voltage $V_{OUT} = 33.14$ V at load ($I_{OUT} = 1$ mA), rise time of the output $t_R = 17.37$ ms, and ripple voltage of the output $V_R = 7.24$ mV (p-p).

The question of an optimal value of the clock frequency is very important in a relationship with the above-mentioned influence of the shoot-through current. The capacitance C = 2.2 μF and the frequency f = 33 kHz calculated from (7) and (8) were used as a reference design. Both these parameters were proportionally changed to values: C = 1 μF and f = 73.3 kHz, C = 4.7 μF and f = 15.6 kHz, C = 10 μF and f = 7.33 kHz. The simulation results are presented in Fig.7.

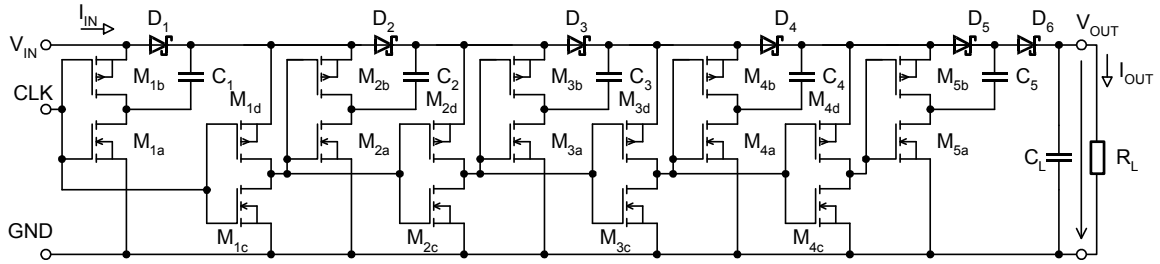

Fig.6. Schematic diagram of the 5-stage FCP.

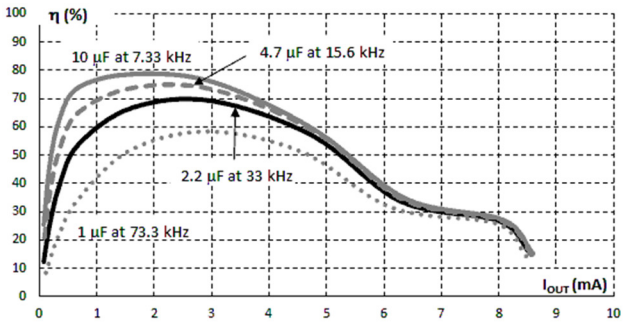

Fig.7. Efficiency vs. output current for various values of capacitance of capacitors.

The increase of the clock frequency by two times approx. to f = 73.3 kHz causes a significant decrease of the efficiency at low values of the output current, whereas, the clock frequency decrease causes the efficiency boosting. The reference design (C = 2.2 µF and f = 33 kHz) was used as a compromise between the efficiency and capacitors values.

3. MEASUREMENT PROCEDURE AND RESULTS

Proposed 11-stage Dickson charge pump (see Fig.4.) and 5-stage Fibonacci charge pump (see Fig.6.) were realised from discrete devices that are listed in Table 1. The realised PCBs contain five terminals for connecting the input voltage V_{IN}, output voltage V_{OUT}, GND (ground), and the clock signal (see Fig.8.).

Fig.8. shows PCB samples of the first realisation of the proposed DCP and FCP. These samples were realised as single-sided PCBs. The second realisation of DCP and FCP were implemented as double-sided PCBs. The 11-stage DCP had dimensions 42.55 × 11.91 mm. The 5-stage FCP had dimensions 40.32 × 12.38 mm.

The key parameters of both charge pumps were measured by the circuit according to Fig.9. The ammeters A_1 and A_2 measure the input and output currents. The input current corresponds to the consumed current, and the output current corresponds to the current of a load. The voltmeters V_1 and V_2 measure the input and output voltage. The used voltmeters had input resistances 20 MΩ.

We compensated a voltage drop of the ammeter A_1 in the time of the measurement. Thus, the input voltage was regulated to value V_{IN} = 3 V accurately. The clock generator produced a square wave signal with frequency 33 kHz and voltage swing 0 to 3 V.

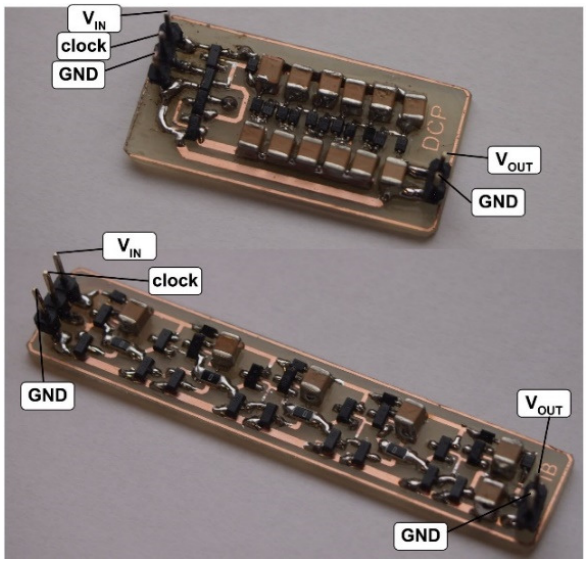

Fig.8. Photography of samples of the first realisation of DCP (top) and FCP (bottom).

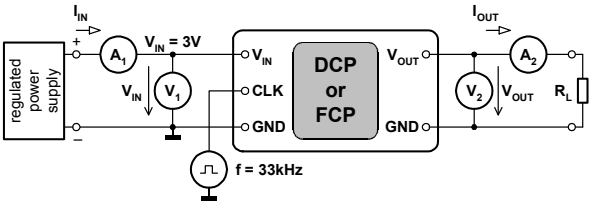

Fig.9. Schematic diagram of the measured circuit.

The FCP propagation delay from the clock input to the clock output of the last inverter (M_{4c}, M_{4d}, see Fig.6.) was 76 ns by the maximum. Thus, the measured value of the propagation delay is sufficiently small in comparison with the clock signal period (for frequency 33 kHz we get period 30 µs approx.).

A. Output voltage vs. output current

The output voltage vs. output current characteristic is a relationship between the output voltage and the corresponding output current. The measured and simulated characteristics for the 11-order DCP and the 5-order FCP are shown in Fig.10.

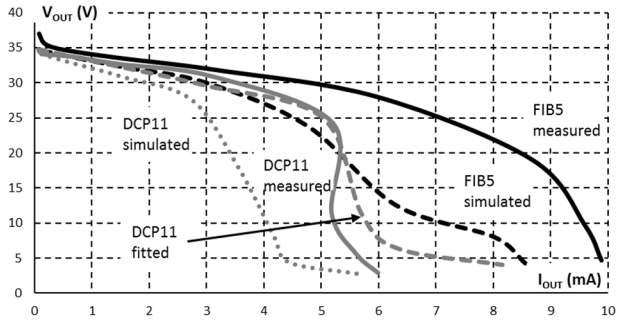

Fig.10. Output voltage vs. output current.

The measured characteristic of the 11-order DCP contains a region with negative differential resistance around I_{OUT} = 5.3 mA (Fig.10., Fig.11., Fig.12., Fig.14., Fig.16.). This effect is caused by simplified construction of DCP clock drivers (see Fig.4.). The used clock drivers (inverters) are loaded by a relatively high capacitance and have not enough driving capacity. Moreover, the first inverter M_1, M_2 drives the second inverter M_3, M_4. The total load of the first inverter caused a significant increase of rising and falling edges of clka and decrease of the clka magnitude. Thus, the second inverter is not optimally driven. The described effect dominates especially at a higher value of the output current. Notice that due to the usage of discrete components, there was a slightly limited choice among MOS transistors. The inverters in the FIB pump are even a bit stronger than what is needed. On the other hand, adding more buffers to the DCP would penalize the efficiency of DCP – therefore we keep the DCP pump for simplicity as it is.

The difference between simulated and measured results is relatively high. This effect is predominantly caused by the threshold voltage variability of used transistors and diodes. The used transistors and diodes have a lower value of the threshold voltage than the value defined in simulated models.

The output series resistance R_S for the output current I_{OUT} = 1 mA can be calculated by [8]

$$R_S = \frac{\Delta V_{OUT}}{\Delta I_{OUT}} = \frac{V_{OUT2} - V_{OUT1}}{I_{OUT1} - I_{OUT2}}, \tag{10}$$

where R_S is the output series resistance, V_{OUT1} is the output voltage at the output current I_{OUT1}, V_{OUT2} is the output voltage at the output current I_{OUT2}.

For the 11-stage DCP were measured values: V_{OUT1} = 33.7 V at I_{OUT1} = 0.714 mA, V_{OUT2} = 32.7 V at I_{OUT2} = 1.485 mA and calculated R_S = 1.30 kΩ. For the 5-stage FCP were measured values: V_{OUT1} = 34.4 V at I_{OUT1} = 0.729 mA, V_{OUT2} = 33.5 V at I_{OUT} = 1.52 mA and calculated R_S = 1.14 kΩ.

B. Efficiency vs. output current

The efficiency of both charge pumps is calculated as the average output power to average input power ratio. The average value of the power is defined by [8]

$$P = \frac{1}{T}\int_0^T v(t) \cdot i(t) \cdot dt, \tag{11}$$

where P is the average value of the power, T is the period, $v(t)$, $i(t)$ are the voltage and current.

The used ammeters measure the average value of a current [9], and the input and output voltage in the steady-state are close to DC. Thus, the calculation of power can be simplified to form (12).

$$P = \frac{1}{T}\int_0^T v(t) \cdot i(t) \cdot dt = \frac{V}{T}\int_0^T i(t) \cdot dt = V \cdot I, \tag{12}$$

where V is the DC voltage, I is the average value of the current.

We calculated the efficiency by (13) from measured values of the ammeters and voltmeters.

$$\eta = \frac{P_{OUT}}{P_{IN}} \cdot 100\,\% = \frac{V_{OUT} \cdot I_{OUT}}{V_{IN} \cdot I_{IN}} \cdot 100\,\%, \tag{13}$$

where η is the efficiency, P_{OUT}, P_{IN} are the average values of the output power and consumed power on input, V_{OUT}, I_{OUT}, V_{IN}, I_{IN} are the measured values of voltages and currents.

The measured and simulated characteristics $\eta = f(I_{OUT})$ for the 11-order DCP and the 5-order FCP are shown in Fig.11.

The difference between simulated and measured results was caused by the threshold voltage variability of used transistors and diodes, again. The measured efficiency of FCP is higher than DCP over the all observed range of the output current.

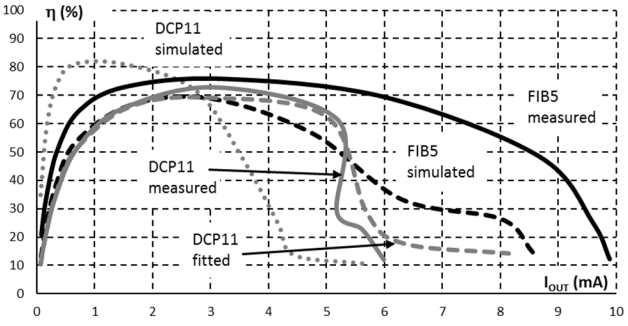

Fig.11. Efficiency vs. output current.

C. Output voltage and efficiency vs. output current for various input voltages

The line regulation is an important characteristic of a charge pump. This characteristic corresponds to decreasing voltage of a system powered by batteries that are gradually discharged. The input voltage V_{IN} was regulated to values 2.8 V, 2.9 V, and 3 V accurately. Simultaneously, the amplitude of clock was set to the same value as the input voltage. The load was set to R_L = 30 kΩ.

The resulting characteristics from Fig.12. show that the output voltage of DCP is higher than the required value V_{OUT} = 30 V at load R_L = 30 kΩ. These values are 31.1 V, 31.7 V, 32.9 V.

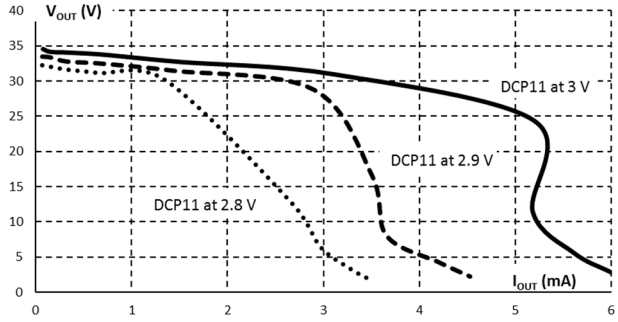

Fig.12. DCP output voltage vs. output current for various input voltages.

The resulting characteristics from Fig.13. show that the output voltage of FCP is higher than the required value V_{OUT} = 30 V at load R_L = 30 kΩ. These values are 31.2 V, 32.3 V, 33.8 V.

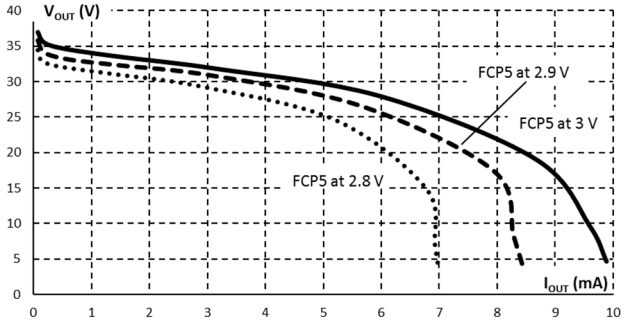

Fig.13. FCP output voltage vs. output current for various input voltages.

Secondly, the influence of the input voltage to the efficiency was measured, see Fig.14. and Fig.15.

The resulting efficiency characteristics of DCP from Fig.14. show that the efficiency for lower values of the output current is independent of the value of the input voltage. These characteristics are very similar up to the output current I_{OUT} = 1.4 mA.

The resulting efficiency characteristics of FCP from Fig.15. show that the efficiency for lower values of the output current is independent of the value of the input voltage. These characteristics are very similar up to the output current I_{OUT} = 3 mA.

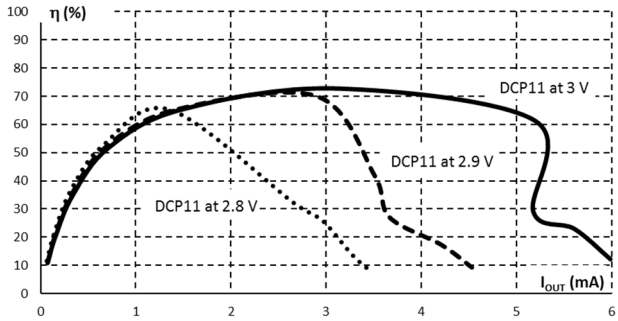

Fig.14. DCP efficiency vs. output current for various input voltages.

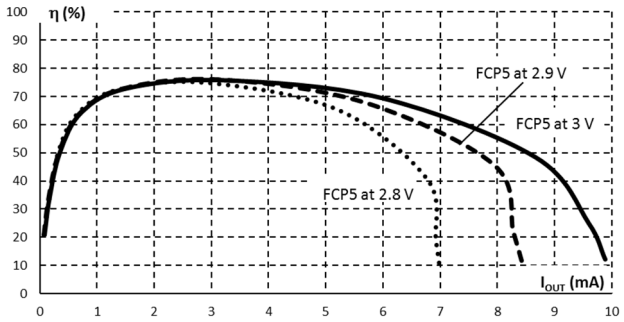

Fig.15. FCP efficiency vs. output current for various input voltages.

D. Output voltage and efficiency vs. output current for various clock frequencies

The clock frequency was subsequently set to values 10 kHz, 33 kHz, and 100 kHz.

The 11-stage DCP produced the output voltage at load R_L = 30 kΩ, V_{OUT} = 32.6 V, 32.9 V, 33.1 V for clock frequency 10 kHz, 33 kHz, and 100 kHz. Thus, the DCP conforms to the required value of the output voltage V_{OUT} = 30 V at the output current I_{OUT} = 1 mA.

The 5-stage FCP produced the output voltage at load R_L = 30 kΩ, V_{OUT} = 27.2 V, 33.8 V, 30.1 V for clock frequency 10 kHz, 33 kHz, and 100 kHz. Thus, the FCP conforms to the required value of the output voltage V_{OUT} = 30 V at the output current I_{OUT} = 1 mA except clock frequency 10 kHz.

The resulting efficiency characteristics of DCP from Fig.16. show that the efficiency for lower values of the output current is independent of the value of the clock frequency. These characteristics are very similar up to the output current I_{OUT} = 2 mA.

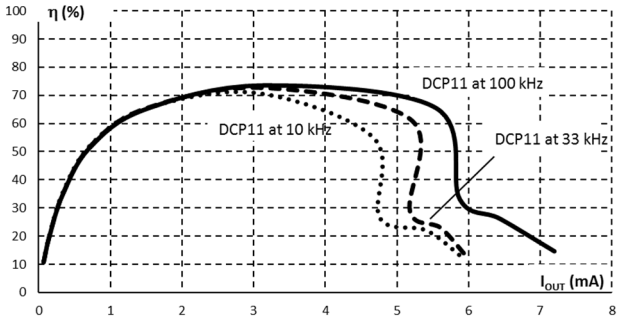

Fig.16. DCP efficiency vs. output current for various clock frequencies.

The resulting efficiency characteristics of FCP from Fig.17. show that the value of optimal clock frequency is 33 kHz. The efficiency is strongly dependent on frequency. At the lower frequencies, the FCP generates a lower output voltage. Thus, the efficiency has a lower value too (13). At the higher frequencies, the cross current of internal FCP inverters is increased. Thus, the input consumed current is increased too. The result is a lower value of the efficiency (13).

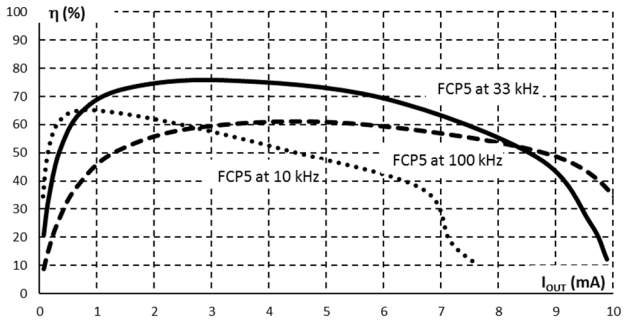

Fig.17. FCP efficiency vs. output current for various clock frequencies.

E. Rise time measurement

The ramp of the output voltage was recorded by a digital oscilloscope in the arrangement depicted in Fig.18. The channel 1 was connected to the input and used as a synchronization source. The channel 2 was used for scanning of the output. The oscilloscope was configured for triggering by channel 1 and for a single shot. The oscilloscope recorded the ramp of the output voltage after closing the switch S.

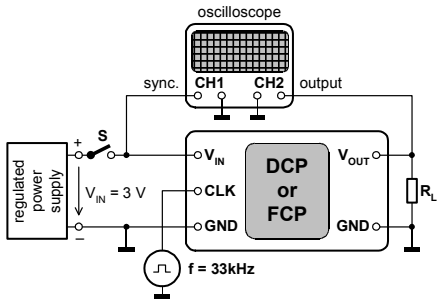

Fig.18. Schematic diagram of the CUT test bench.

These oscillograms were recorded for the load resistance R_L = 30 kΩ. The used value of the load resistance implies the required minimal output voltage V_{OUT} = 30 V at current I_{OUT} = 1 mA.

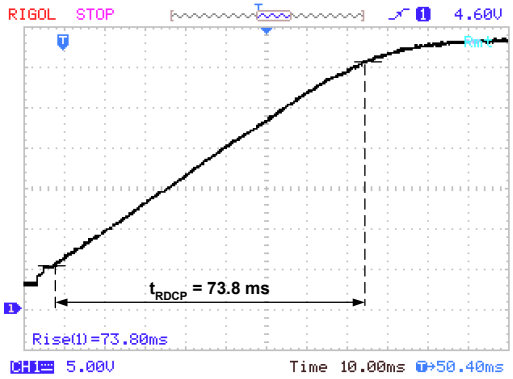

Fig.19. Oscillogram of the ramp of the DCP output at load R_L = 30 kΩ.

Both charge pumps produced the steady-state output voltage higher than the required minimal value. The 11-stage DCP had the rise time t_{RDCP} = 73.8 ms and the 5-stage FCP had the rise time t_{RFCP} = 14.7 ms.

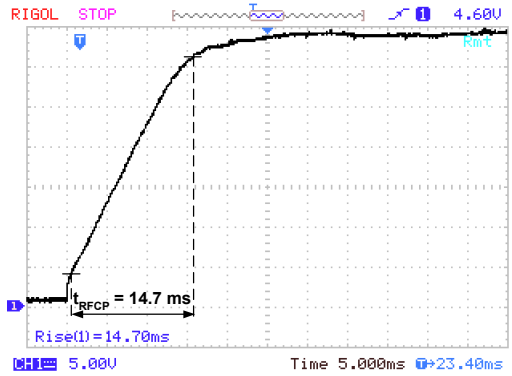

Fig.20. Oscillogram of the ramp of the FCP output at load R_L = 30 kΩ.

4. RESULTS

The key parameters of the 11-stage Dickson charge pump and 5-stage Fibonacci charge pump are summarised and compared in Table 2.

Table 2. Comparison of simulated and measured results (V_{IN} = 3 V, f = 33 kHz, R_L = 30 kΩ).

Parameter	DCP11 sim	DCP11 meas	FCP5 sim	FCP5 meas
V_{OUT}	31.7 V	32.9 V	32.6	33.8 V
η	81 %	60 %	61 %	69 %
R_S	2.24 kΩ	1.30 kΩ	1.58 kΩ	1.14 kΩ
t_R	45.3 ms	73.8 ms	17.9 ms	14.7 ms
V_{Rp-p}	7 mV	11 mV	7 mV	10 mV
N_C	12	12	6	6
N_D	12	12	6	6
N_T	4	4	18	18

DCP11sim, *DCP11meas* mean the results from simulation or measurement of DCP, *FCP5sim*, *FCP5meas* mean the results from simulation or measurement of FCP, V_{IN}, V_{OUT} are the input and output voltages, f is the clock frequency, R_L is the load resistance, η is the efficiency, R_S is the output series resistance for output current 1 mA, t_R is the rise time of the output, V_{Rp-p} is the peek-to-peek ripple voltage of the output, and N_C, N_D, N_T are the required number of capacitors, diodes, and transistors.

Compared to DCP, the extra costs on FIB could potentially allow making a high-performance CP eliminating effectively the number of capacitors in the design. Notice that low loss capacitors could be expensive components. This fact, together with a limited number of low loss capacitors, tends to achieve much higher efficiency.

5. CONCLUSION

In this article, a circuit solution compact implementation of the Fibonacci pump was presented, including the detailed design procedure considering an improved clock buffer scheme.

The key parameters of the Dickson charge pump and the Fibonacci charge pump were simulated and then measured. Some differences between simulations a measurement are evoked by a variability of parameters of used devices. E.g. variability of the threshold voltage of used transistors or forward drop voltage of used diodes has a strong effect on many observed parameters. The diode and transistor threshold voltage spread was taken into account by post-fitting of the device model parameters. The reverse bias saturation current and the series resistance of the diode model were changed to values $I_S = 600$ μA and $R_S = 0.1$ Ω (original values: $I_S = 2.831$ μA and $R_S = 0.1975$ Ω). The threshold voltage of NMOS transistor model was changed to value $V_{T0} = 1.5$ V (original value: $V_{T0} = 1.6$ V). The threshold voltage of PMOS transistor model was changed to value $V_{T0} = -1.5$ V (original value: $V_{T0} = -2.1$ V). Fig.10. and Fig.11. show the results from simulations after fitting models of used transistors and diodes for the 11-stage Dickson charge pump. Now, the fitted and measured results are relatively close.

The realised Fibonacci charge pump has higher values of the efficiency and output voltage than the Dickson charge pump. The Fibonacci charge pump is suitable for a higher value of the voltage gain and requires more transistors than the Dickson charge pump. But the number of used diodes and capacitors for the Fibonacci charge pump is lower than for the Dickson charge pump. The Fibonacci pump concept is especially attractive to the pumps realised by discrete components, as the voltage gain for an intermediate number of stages is high. On the other hand, this charge pump type is not so suitable for integration into ASICs, because its sensitivity for on-chip parasitics is higher than e.g. for the Dickson-based architectures.

The measured parameters verify a possibility of the realisation of the Dickson and the Fibonacci charge pump in its discrete form. In the next period, the issue of the clock drivers for DCP will be resolved. The driving capability will be increased by splitting capacitors into sections. Each section will be driven by a separate clock driver. The FCP will be then extended by an auxiliary current limiter for the inverter in each stage. Other types of transistors with a lower threshold voltage will be finally used.

ACKNOWLEDGMENT

This work has been supported by the grant No. SGS17/183/OHK3/3T/13 of the CTU in Prague.

REFERENCES

[1] Dickson, J.F. (1976). On-chip high-voltage generation in NMOS integrated circuits using an improved voltage multiplier technique. *IEEE Journal of Solid-State Circuits*, 11 (3), 374-378.

[2] Pan, F., Samaddar, T. (2006). *Charge Pump Circuit Design*. McGraw-Hill Education.

[3] Seeman, M.D., Sanders, S.R. (2008). Analysis and optimization of switched-capacitor DC–DC converters. *IEEE Transactions on Power Electronics*, 23 (2), 841-851.

[4] Ueno, F., Inoue, T., Oota, I., Harada, I. (1991). Emergency power supply for small computer systems. In *IEEE International Symposium on Circuits and Systems*, 11-14 June 1991. IEEE, 1065-1068.

[5] Tanzawa, T. (2016). Innovation of switched-capacitor voltage multiplier: Part 1: A brief history. *IEEE Solid-State Circuits Magazine*, 8 (1), 51-59.

[6] Allasasmeh, Y., Gregori, S. (2009). A performance comparison of dickson and fibonacci charge pumps. In *European Conference on Circuit Theory and Design*, 23-27 August 2009. IEEE, 599-602.

[7] Matousek, D., Hospodka, J., Subrt, O. (2016). Efficiency of innovative charge pump versus clock frequency and MOSFETs sizes. *Measurement Science Review*, 16 (5), 260-265.

[8] Mayergoyz, I., Lawson, W. (2012). *Basic Electric Circuit Theory, 2nd Edition.* Academic Press.

[9] Tumanski, S. (2006). *Principles of Electrical Measurement*. CRC Press.

Received February 20, 2017.
Accepted April 28, 2017.

MEASUREMENT SCIENCE REVIEW, **17**, (2017), No. 3, 108-116

DE GRUYTER OPEN

MEASUREMENT SCIENCE REVIEW

Journal homepage: http://www.degruyter.com/view/j/msr

ISSN 1335-8871

Application of Monte Carlo Method for Evaluation of Uncertainties of ITS-90 by Standard Platinum Resistance Thermometer

Rudolf Palenčár, Peter Sopkuliak, Jakub Palenčár, Stanislav Ďuriš, Emil Suroviak, Martin Halaj

Faculty of Mechanical Engineering, Institute of Automation, Measurements and Applied Informatics, Slovak Technical University, 81243 Bratislava, Nám. Slobody 17, Slovak Republic, email: rudolf.palencar@stuba.sk, xsopkuliak@is.stuba.sk

Evaluation of uncertainties of the temperature measurement by standard platinum resistance thermometer calibrated at the defining fixed points according to ITS-90 is a problem that can be solved in different ways. The paper presents a procedure based on the propagation of distributions using the Monte Carlo method. The procedure employs generation of pseudo-random numbers for the input variables of resistances at the defining fixed points, supposing the multivariate Gaussian distribution for input quantities. This allows taking into account the correlations among resistances at the defining fixed points. Assumption of Gaussian probability density function is acceptable, with respect to the several sources of uncertainties of resistances. In the case of uncorrelated resistances at the defining fixed points, the method is applicable to any probability density function. Validation of the law of propagation of uncertainty using the Monte Carlo method is presented on the example of specific data for 25 Ω standard platinum resistance thermometer in the temperature range from 0 to 660 °C. Using this example, we demonstrate suitability of the method by validation of its results.

Keywords: The law of propagation of uncertainty, Monte Carlo method, the International Temperature Scale of 1990 (ITS-90), Standard Platinum Resistance Thermometer.

1. INTRODUCTION

This paper presents a method based on the propagation of distributions by Monte Carlo method (MCM). The procedure is based on the generation of pseudo-random numbers of input variables of multi-dimensional distribution. Multi-dimensional distribution is used because it takes into account correlation among the Standard Platinum Resistance Thermometer (SPRT) resistances from calibration as well as the SPRT resistances in temperature measurement. Generating input variables only from the one-dimensional distribution is sufficient for uncorrelated resistances.

In our case it is necessary to identify the probability distributions of input quantities and relevant multivariate distribution function for the case of correlated input quantities. We can assume normal distribution for all input SPRT resistances and therefore multivariate normal distribution for correlated resistances. This assumption is based on the central limit theorem, because several sources of uncertainties are present at the measurement: e.g. self-heating effect of the SPRT, chemical impurities of the substance in defining fixed points (DFPs), immersion effect of the SPRT, hydrostatic-head effect, effect of gas pressure in DFPs, choice of fixed point value from plateau isotopic variations, residual gas pressure in triple point of water (TPW) cell, non-linearity of the resistance bridge, changes of resistances of standard resistor initiated by changes of its temperature, uncertainty of calibration of resistance standard, etc.

The aim of this study is

a) presentation of the MCM for uncertainty evaluation of the international temperature scale ITS-90 by using SPRT calibrated at DFPs;
b) validation of the process by using the law of propagation of uncertainty according to the Guide to the Expression of Uncertainty in Measurement GUM [1] for specific conditions.

The procedure was designed to take into account correlations among the SPRT resistances obtained from calibration at the DFPs and those obtained in temperature measurements.

Influences such as fluctuations, drifts, temperature gradients and similar are not analyzed. Also, the uncertainties caused by non-uniqueness and consistency sub-ranges are not included. These questions are presented e.g. in [2]. The case study focuses on determination of uncertainties related to realization of international temperature scale by using SPRT calibrated in the range from 0 °C to the freezing point of aluminum (660.323 °C).

DOI: 10.1515/msr-2017-0014

2. CURRENT STATUS OF THE ISSUE

Most published papers employ an approach based on the law of propagation of uncertainty GUM and its supplements [3] and [4], fewer papers are based on the orthogonal polynomials. The overview of the approaches is presented in [5], [6]. Matrix interpretation of GUM method is described in [7]. Specific approaches dealing with density functions and confidence intervals are mentioned in [8], [9]. The effect of covariance on uncertainty when realizing ITS-90 temperature scale is discussed in [6]. Uncertainty determination by MCM is discussed e.g. in [10], [11]. Propagation of distributions using MCM, based on Supplement 1 [3] to the GUM [1], occurs only in a few isolated cases e.g. [12], [13]. In general, authors predict uncorrelated resistances among the defining fixed point (DFP) and neglect the influence of correlations. Omitting the correlations among resistances in individual DFPs does not always correspond to reality and they can have a significant impact, see [6]. Progressive Bayesian analysis introduces another point of view. Simple measurement model is presented in [14].

Deviation equations and ratios of resistance belong to a non-linear model, defined by ITS-90. The non-linear model can generate some doubts about the adequacy of its linearization by expansion in Taylor series of the first order. Also, complications with the determination of the sensitivity coefficients may occur. In such cases, Monte Carlo method, based on propagation of distributions, is preferably used. Papers on the Monte Carlo method, which follows the recommendations and procedures listed in [3], appear sporadically. The basic principle of both methods is shown in Fig.1.

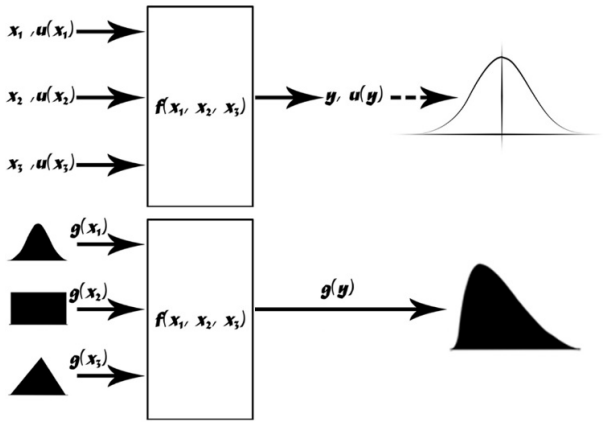

Fig.1. The law of propagation of uncertainty (up) and the law of propagation of distribution (down).

3. THEORETICAL BASES OF ITS-90

International Temperature scale of 1990 defines the temperature from the inverse function

$$T = f(W_r) \tag{1}$$

Corresponding sub-ranges of the ITS-90 from 0 °C for the functions (1) are stated in [15]. Function W_r is given by

$$W_r = W - \sum_{i=1}^{N} a_i f_i(W) \tag{2}$$

where

$$W = \frac{R}{R_{\mathrm{TPW}}} \tag{3}$$

while

R is the SPRT resistance at temperature T and R_{TPW} is the SPRT resistance at TPW

$f_i(W)$ are functions of the individual sub-ranges, see [15].

a_i are coefficients of deviation function from the calibration of SPRT at DFPs and

$a_i = g_1\,(R_{\mathrm{TPW1}}, \dots, R_{\mathrm{TPW}N}, R_{\mathrm{DFP1}}, \dots, R_{DFPN})$ or $a_i = g_2\,(W_{\mathrm{DFP1}},\ W_{\mathrm{DFP2}}, \dots, W_{DFPN})$.

Matrix notation for the calculation of the coefficients of deviation function can be used. If the relationship (2) is applied to N fixed points, then

$$\begin{pmatrix} \Delta W_{\mathrm{DFP1}} \\ \vdots \\ \Delta W_{\mathrm{DFP}N} \end{pmatrix} = \begin{pmatrix} f_1(W_{\mathrm{DFP1}}) & \cdots & f_N(W_{\mathrm{DFP1}}) \\ \vdots & \ddots & \vdots \\ f_1(W_{\mathrm{DFP}N}) & \cdots & f_N(W_{\mathrm{DFP}N}) \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix} \tag{4}$$

where $\Delta W_{\mathrm{DFP}i} = W_{\mathrm{r,DFP}i} - W_{\mathrm{DFP}i}$, $W_{\mathrm{DFP}i}$ are resistance ratios for corresponding DFPi and $W_{\mathrm{r,DFP}i}$ are defined in [15].

Equation (4) is written in the form

$$\Delta \boldsymbol{W}_{\mathrm{DFP}} = \boldsymbol{M}_{\mathrm{DFP}}\, \boldsymbol{a} \tag{5}$$

As $\boldsymbol{M}_{\mathrm{DFP}}^{-1}$ exists, coefficients of deviation function are given by

$$\boldsymbol{a} = \boldsymbol{M}_{\mathrm{DFP}}^{-1}\, \Delta \boldsymbol{W}_{\mathrm{DFP}} \tag{6}$$

Then the equation (2) can be rewritten as follows

$$W_r = W + \boldsymbol{a}^T \boldsymbol{f}(\boldsymbol{W}). \tag{7}$$

4. APPLICATION OF MONTE CARLO METHOD

A. The input data and basic relations

Input data for Monte Carlo method are represented by SPRT resistances at defining fixed points, obtained from calibration and SPRT resistances obtained from temperature measurement. Thus, input quantities can be written as vector of dimension $2N + 2$ where

$$\boldsymbol{R} = (\,R_{\mathrm{TPW1}}, \dots, R_{\mathrm{TPW}N}, R_{\mathrm{DFP1}}, \dots, R_{\mathrm{DFP}N},\ R_{\mathrm{TPW}}, R)^{\mathrm{T}}. \tag{8}$$

Covariance matrix and vector of input quantities will be in form $(2 + 2\mathrm{N})\times(2 + 2N)$. The resistances of SPRT are determined on the basis of SPRT calibration at DFPs, i.e. vector

$$\boldsymbol{R}_{\mathrm{cal}} = (R_{\mathrm{TPW1}}, \dots, R_{\mathrm{TPW}N}, R_{\mathrm{DFP1}}, \dots, R_{\mathrm{DFP}N})^{\mathrm{T}}. \tag{9}$$

SPRT resistances R corresponding to measured temperatures T, as well as SPRT resistances at triple point of

water R_{TPW}, their covariances and uncertainties are determined in phase of temperature measurement. It implies the vector $\boldsymbol{R}_{\mathrm{meas}} = (R, R_{\mathrm{TPW}})^{\mathrm{T}}$ and its covariance matrix. Beside that it is necessary to consider the covariances among SPRT resistances obtained from calibration and from temperature measurement. Then we can write the covariance matrix of the vector (8) in the form

$$\boldsymbol{V}_{\mathrm{R}} = \begin{pmatrix} \boldsymbol{V}_{\boldsymbol{R}_{\mathrm{meas}}} & \boldsymbol{V}_{\boldsymbol{R}_{\mathrm{meas}},\boldsymbol{R}_{\mathrm{cal}}} \\ \boldsymbol{V}_{\boldsymbol{R}_{\mathrm{meas}},\boldsymbol{R}_{\mathrm{cal}}} & \boldsymbol{V}_{\boldsymbol{R}_{\mathrm{cal}}} \end{pmatrix}. \tag{10}$$

Covariance matrix $\boldsymbol{V}_{\boldsymbol{R}_{\mathrm{meas}}}$ expresses the uncertainties of SPRT resistances and covariances among them for temperature measurements in equation (10). Covariance matrix $\boldsymbol{V}_{\boldsymbol{R}_{\mathrm{cal}}}$ expresses uncertainties of SPRT resistances and covariances among them for calibration of SPRT at DFPs. Matrix $\boldsymbol{V}_{\boldsymbol{R}_{\mathrm{meas}},\boldsymbol{R}_{\mathrm{cal}}}$ expresses the covariances among the SPRT resistances from measurement and resistances from calibration. In case of in-house temperature measurement, i.e. SPRT resistance value at TPW is used from calibration, matrix $\boldsymbol{V}_{\boldsymbol{R}_{\mathrm{meas}},\boldsymbol{R}_{\mathrm{cal}}}$ has nonzero values. If we suppose that SPRT resistances during temperature measurement were obtained under the same conditions as they were in calibration process, then there could be also covariances between SPRT resistances for temperature measurement and SPRT resistances from calibration.

Usually, in practice, the covariances among SPRT resistances are not considered, excluding those at TPW. Two cases of temperature measurement are considered here, a) the use of calibrated SPRT in the laboratory (in-house), b) the use of calibrated SPRT outside the laboratory.

B. Procedure of calculation

Fig.2. schematically shows the process of calculation the temperature and its standard uncertainty by using Monte Carlo method [3].

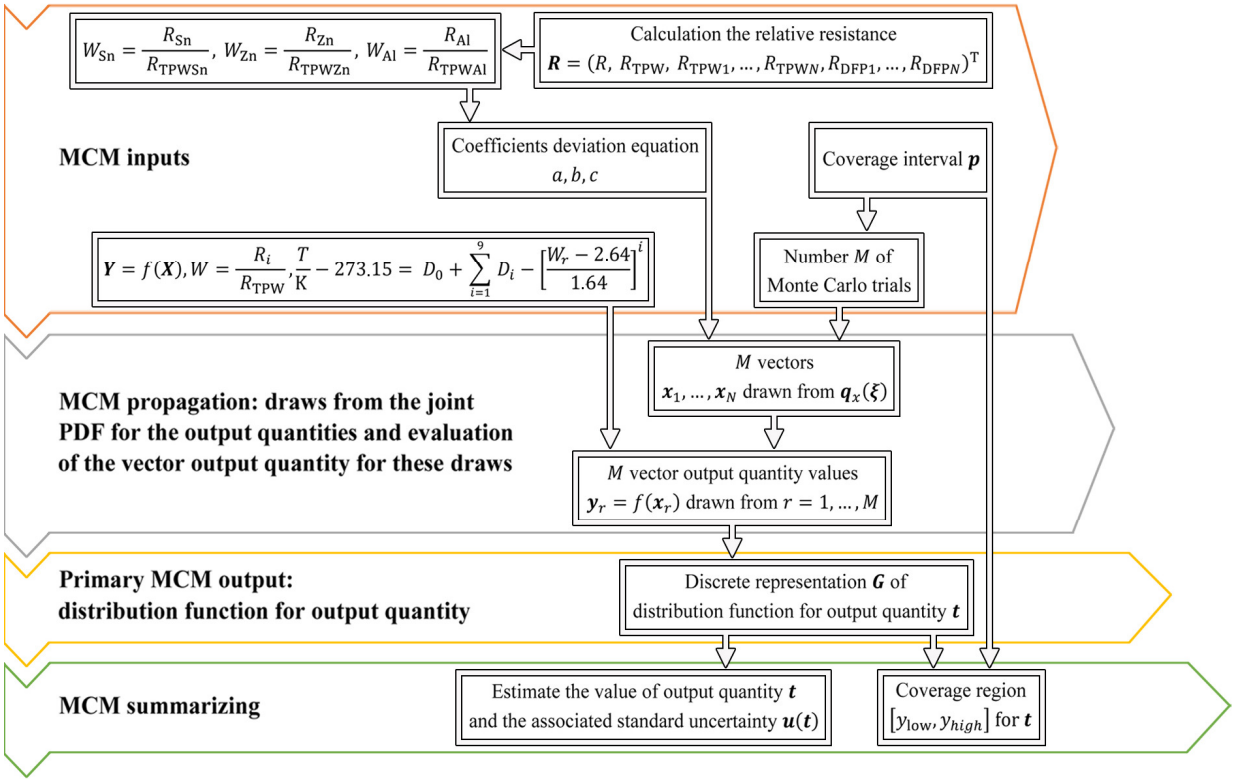

Fig.2. Computing phase of calibration using Monte Carlo method [3].

C. Used software and pseudo-random numbers generation

When selecting a suitable programming software environment, it was crucial to create an application which would be easy to use and portable. It was also necessary to consider the efficiency of the calculation of final application and the way of implementation of Mersenne Twister generator. For these reasons, Microsoft Visual Basic.NET programming environment was chosen and 32-bit version of operating system has been used due to direct compatibility with newer 64-bit operating systems.

Visual Basic does not integrate the generation of pseudo-random numbers with Mersenne Twister (MT) generator which is currently the best rated algorithm which has undergone a large number of experiments for testing pseudo-random numbers. Therefore, the final algorithm uses the original MT source code translated to VB.NET framework. Since MT generator generates numbers from uniform distribution, it was necessary to use the Box-Muller transformation method. This method allows transforming uniformly distributed random variables to the Gaussian

distribution. The application of the Box-Muller transformation method can be simplified by utilizing the fact that the MT generator allows direct generation of uniformly distributed numbers over the interval [-1, 1]. Input variables x_1 and x_2 take values from listed interval and subsequently enter formula $w = x_1^2 + x_2^2$. This loop repeats until the condition $w \leq 1$ is true. Afterwards $u = \sqrt{[-2 \cdot \log(w)]/w}$ can be calculated and two independent output pseudo-random variables with normal distribution $N(0,1)$ are obtained, using $y_1 = x_1 \cdot u$ and $y_2 = x_2 \cdot u$. Normal distribution for any mean value μ and the variance σ can be obtained as $\boldsymbol{\mu} + \boldsymbol{\sigma} \cdot \boldsymbol{z}$ where $\boldsymbol{z}$ is a matrix of randomly generated values from standard normal distribution $N(0,1)$. In our case, normal distribution is assumed for all input quantities and they are correlated in general.

If we want to construct a generator of pseudo-random numbers from a multi-dimensional normal distribution $N(\boldsymbol{\mu}, \boldsymbol{V})$, it is essential to establish dimension n of multidimensional normal distribution, vector of mean values $\boldsymbol{\mu}$ of dimension $n \times 1$, covariance matrix $\boldsymbol{V}$ of dimension $n \times n$ and the number of trials that should be generated. A matrix $\boldsymbol{X}$ of dimension $n \times q$ must be generated as well. We derive $\boldsymbol{R}^{\mathrm{T}}$ from the covariance matrix $\boldsymbol{V}$ by using the Cholesky decomposition $\boldsymbol{V} = \boldsymbol{R}^{\mathrm{T}}\boldsymbol{R}$. We will generate a matrix $\boldsymbol{Z}$ of dimension $n \times q$ from the normal distribution. Then we compute $\boldsymbol{X} = \boldsymbol{\mu}\boldsymbol{1}^{\mathrm{T}} + \boldsymbol{R}^{\mathrm{T}}\boldsymbol{Z}$, where $\boldsymbol{1}$ denotes the unit vector of dimension q (see [3]). The number of Monte Carlo trials M that must be carried out for each sequence h, must be determined for calculation of the estimate of temperature T and its standard uncertainty $u(T)$ by adaptive Monte Carlo method. For the output quantity T we have to consider the reference probability p and the number of significant decimal digits n_{dig} from a standard uncertainty $u(T)$. Number of Monte Carlo trials M increases as $(M \times h)$ by each further sequence of calculation h to stabilize the required statistical output quantities.

5. Evaluation of uncertainty and experimental data

In order to compare the results of both realized cases (calibration in-house and outside the laboratory), we employ the evaluation data obtained from the Slovak Institute of Metrology (SMU) – see Table 1. These values will be used as inputs for evaluating of realization of ITS-90 and corresponding uncertainties by Monte Carlo method.

A. Inputs and considered cases

The input data are given in Table 1. We consider the temperature measurement according to ITS-90 first in the calibration laboratory (in-house), then outside the calibration laboratory. For temperature measurement, the same TPW cell is used as was used for realization of temperature scale. In this case, the SPRT resistances at TPW after tin, zinc and aluminum are correlated (for sake of simplicity we consider the correlation coefficient r = 1). SPRT resistance of temperature measurement is considered uncorrelated with the other SPRT resistances. SPRT resistances at DFPs are considered either uncorrelated (cases a), b)), or correlated (cases c), d)). The correlation coefficients among resistances at TPW and DFPs are considered uniformly r = 0.4 while correlation coefficients among the resistances at DFPs are uniformly considered as r = 0.3.

Case a) SPRT is used in-house, resistances at DFPs are uncorrelated

$$\boldsymbol{R}_R = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\boldsymbol{V}_R = 10^{-10} \begin{pmatrix} 1.61 & 1.61 & 1.61 & 0 & 0 & 0 & 1.61 & 0 \\ 1.61 & 1.61 & 1.61 & 0 & 0 & 0 & 1.61 & 0 \\ 1.61 & 1.61 & 1.61 & 0 & 0 & 0 & 1.61 & 0 \\ 0 & 0 & 0 & 14.8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 24.8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 39.9 & 0 & 0 \\ 1.61 & 1.61 & 1.61 & 0 & 0 & 0 & 1.61 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4.00 \end{pmatrix}$$

Case b) SPRT is used outside calibration laboratory, resistances at DFPs are uncorrelated

$$\boldsymbol{R}_R = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\boldsymbol{V}_R = 10^{-10} \begin{pmatrix} 1.61 & 1.61 & 1.61 & 0 & 0 & 0 & 0 & 0 \\ 1.61 & 1.61 & 1.61 & 0 & 0 & 0 & 0 & 0 \\ 1.61 & 1.61 & 1.61 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 14.8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 24.8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 39.9 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1.61 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4.00 \end{pmatrix}$$

Case c) SPRT is used in-house, resistances at DFPs are correlated

$$\boldsymbol{R}_R = \begin{pmatrix} 1 & 1 & 1 & 0.4 & 0.4 & 0.4 & 1 & 0 \\ 1 & 1 & 1 & 0.4 & 0.4 & 0.4 & 1 & 0 \\ 1 & 1 & 1 & 0.4 & 0.4 & 0.4 & 1 & 0 \\ 0.4 & 0.4 & 0.4 & 1 & 0.3 & 0.3 & 0.4 & 0 \\ 0.4 & 0.4 & 0.4 & 0.3 & 1 & 0.3 & 0.4 & 0 \\ 0.4 & 0.4 & 0.4 & 0.3 & 0.3 & 1 & 0.4 & 0 \\ 1 & 1 & 1 & 0.4 & 0.4 & 0.4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\boldsymbol{V}_R = 10^{-10} \begin{pmatrix} 1.61 & 1.61 & 1.61 & 1.96 & 2.53 & 3.21 & 1.61 & 0 \\ 1.61 & 1.61 & 1.61 & 1.96 & 2.53 & 3.21 & 1.61 & 0 \\ 1.61 & 1.61 & 1.61 & 1.96 & 2.53 & 3.21 & 1.61 & 0 \\ 1.96 & 1.96 & 1.96 & 14.8 & 5.75 & 7.30 & 1.96 & 0 \\ 2.53 & 2.53 & 2.53 & 5.75 & 24.8 & 9.44 & 2.53 & 0 \\ 3.21 & 3.21 & 3.21 & 7.30 & 9.44 & 39.9 & 3.21 & 0 \\ 1.61 & 1.61 & 1.61 & 1.96 & 2.53 & 3.21 & 1.61 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4.00 \end{pmatrix}$$

Case d) SPRT is used outside calibration laboratory, resistances at DFPs are correlated

$$\boldsymbol{R}_R = \begin{pmatrix} 1 & 1 & 1 & 0.4 & 0.4 & 0.4 & 0 & 0 \\ 1 & 1 & 1 & 0.4 & 0.4 & 0.4 & 0 & 0 \\ 1 & 1 & 1 & 0.4 & 0.4 & 0.4 & 0 & 0 \\ 0.4 & 0.4 & 0.4 & 1 & 0.3 & 0.3 & 0 & 0 \\ 0.4 & 0.4 & 0.4 & 0.3 & 1 & 0.3 & 0 & 0 \\ 0.4 & 0.4 & 0.4 & 0.3 & 0.3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\boldsymbol{V}_R = 10^{-10} \begin{pmatrix} 1.61 & 1.61 & 1.61 & 1.96 & 2.53 & 3.21 & 0 & 0 \\ 1.61 & 1.61 & 1.61 & 1.96 & 2.53 & 3.21 & 0 & 0 \\ 1.61 & 1.61 & 1.61 & 1.96 & 2.53 & 3.21 & 0 & 0 \\ 1.96 & 1.96 & 1.96 & 14.8 & 5.75 & 7.30 & 0 & 0 \\ 2.53 & 2.53 & 2.53 & 5.75 & 24.8 & 9.44 & 0 & 0 \\ 3.21 & 3.21 & 3.21 & 7.30 & 9.44 & 39.9 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1.61 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4.00 \end{pmatrix}$$

Whereby $\boldsymbol{V}_R = \boldsymbol{P}_R \boldsymbol{R}_R \boldsymbol{P}_R^{\mathrm{T}}$ and $\boldsymbol{P}_R$ is a diagonal matrix of dimension 8×8 with diagonal elements: $u(R_{\mathrm{TPW\,Sn}}), u(R_{\mathrm{TPW\,Zn}}), u(R_{\mathrm{TPW\,Al}}), u(R_{\mathrm{Sn}}), u(R_{\mathrm{Zn}}), u(R_{\mathrm{Al}}), u(R_{\mathrm{TPW}}), u(R)$.

The results of simulation by Monte Carlo method and GUM are presented in Table 4. The graphical comparison of both methods for 66 calibration points within the range (0 ÷ 660) °C of the ITS-90 is illustrated in Fig.10.

Table 1. Measured values of SPRT resistances in defining fixed point.

Defining fixed point	Resistance (Ω)	Standard uncertainty of resistance (Ω)
Sn	46.9397533	3.85×10^{-5}
Zn	63.7056752	4.98×10^{-5}
Al	83.7191875	6.32×10^{-5}
$\mathrm{TPW_{Sn}}$	24.8002001	1.27×10^{-5}
$\mathrm{TPW_{Zn}}$	24.8001927	1.27×10^{-5}
$\mathrm{TPW_{Al}}$	24.8001872	1.27×10^{-5}

We consider $M = 10^5$ Monte Carlo trials. For the output quantity T we consider the reference probability $p = 0.95$ and the number of significant decimals of the standard uncertainty $n_{\mathrm{dig}} = 2$. Fig.3. shows a histogram of resistance R_{Al}, as given in the example 5.B.

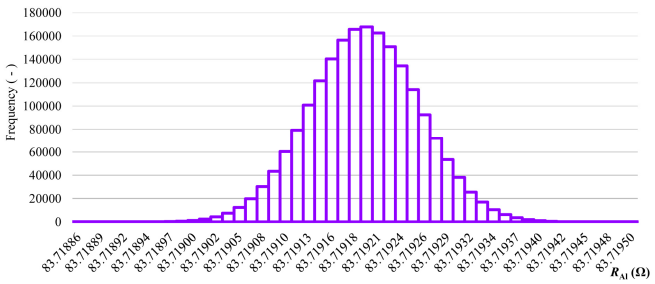

Fig.3. Histogram of input resistance.

The coefficients a, b, c of deviation function can be determined from equation (6), see Fig.5. for their calculation. Histograms of coefficients of deviation function and the coefficient a, presented in Fig.4., have similar shape.

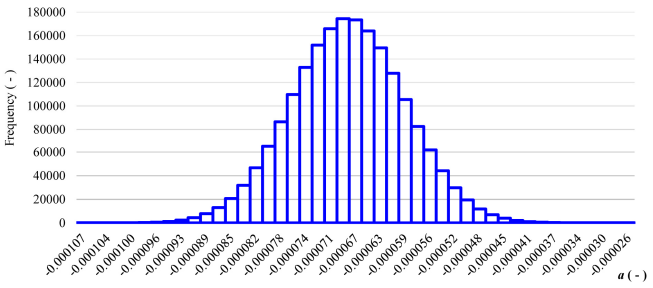

Fig.4. Histogram of the deviation equation for coefficient a.

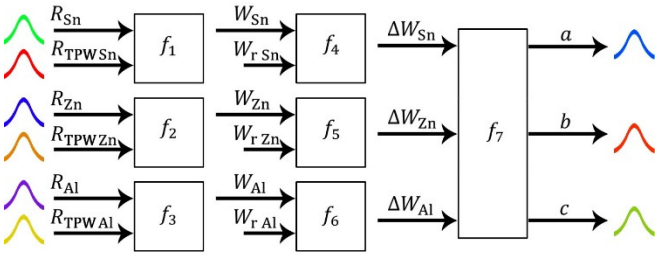

Fig.5. Sub-model for calculation of the coefficients of deviation function.

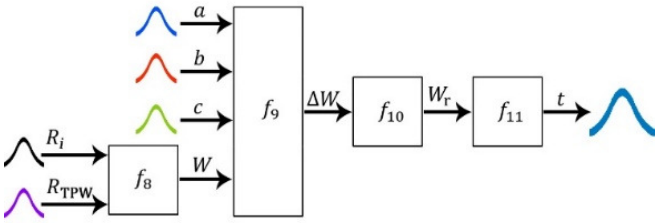

Fig.6. Model to calculate temperature and corresponding standard uncertainty.

Fig.6. shows the calculation procedure for determination of the temperature T, where $f_8 = \frac{R}{R_{\mathrm{TPW}}}$, from equation (2) we can determine $f_9 = a(W-1) + b(W-1)^2 + c(W-1)^3$ and from equation (7) we get $f_{10} = W - a(W-1) + b(W-1)^2 + c(W-1)^3$. Histogram of estimated temperature t is presented in Fig.8.

B. Example of calculating the output characteristics for the case b)

Temperature according to ITS-90 and corresponding standard uncertainty was determined for data in Table 1. and for correlation matrix of resistances for case b, see Table 2.

Table 2. Evaluation of specific resistance using the law of propagation of uncertainty.

R_i (Ω)	t_i (°C)	$u(t_i)$ (°C)
46.55489887	227.75076	5.59864×10^{-4}

Let's use the MCM with $h = 20$ and $M = 10^5$. Based on the generated input values for resistances and using appropriate relationship, we get estimate of the temperature t (°C).

$$t = \begin{bmatrix} 227.7479428 \\ \vdots \\ 227.7533240 \end{bmatrix} \quad (11)$$

The symmetrical reference interval with the specified probability for the estimating output quantity t is obtained from its generated discrete representation (distribution function shown in Fig.7.). To do so, generated values are arranged at non-decreasing sequence, using rules listed in [3]. According to that procedure, we get the following symmetric confidence interval

$$t = [y_{\min}\,; y_{\max}] = [227.74966142\,; 227.75185372] \quad (12)$$

and following narrowest confidence interval

$$t = [y_{\min}\,; y_{\max}] = [227.74966148\,; 227.75185375] \quad (13)$$

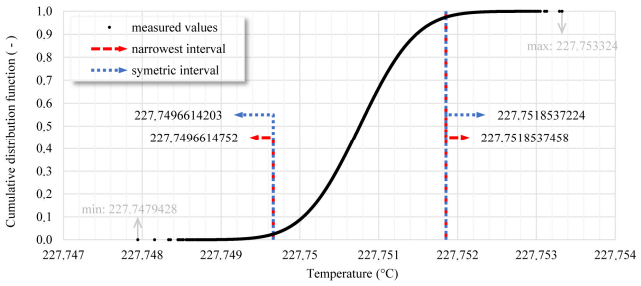

Fig.7. The distribution function of the output temperature.

Statistical characteristics of the resulting estimates are calculated from the partial estimates in each sequence $y^{(h)}$, $u(y^{(h)})$. After calculating the last sequence h, it is possible to calculate the resulting parameters for estimation. In this case

$$\hat{y} = y = \frac{1}{h}\sum_{i=1}^{h} y^{(i)} = 227.7507574\ °C \quad (14)$$

$$s_{\hat{y}} = \sqrt{\frac{1}{h(h-1)}\sum_{i=1}^{h}(y^{(i)} - \hat{y})^2} = 3.39088 \times 10^{-7}\ °C \quad (15)$$

The standard uncertainty

$$\hat{u}(y) = \frac{1}{h}\sum_{i=1}^{h} u(y^{(i)}) = 5.59883 \times 10^{-4} \quad (16)$$

$$s_{\hat{u}(y)} = \sqrt{\frac{1}{h(h-1)}\sum_{i=1}^{h}\left(u(y^{(i)}) - \hat{u}(y)\right)^2} = 2.70938 \times 10^{-7}\ °C \quad (17)$$

The lower end point of the narrowest confidence interval of stabilization criteria

$$\hat{y}_{\min} = \frac{1}{h}\sum_{i=1}^{h} y_{\min}^{(i)} = \frac{1}{20}\left(y_{\min}^{(1)} + y_{\min}^{(2)} + \ldots + y_{\min}^{(20)}\right) = 227.7496603\ °C \quad (18)$$

$$s_{\hat{y}_{\min}} = \sqrt{\frac{1}{h(h-1)}\sum_{i=1}^{h=20}\left(y_{\min}^{(i)} - \hat{y}_{\min}\right)^2} = 2.31921 \times 10^{-6}\ °C \quad (19)$$

The upper end point of the narrowest confidence interval of stabilization criteria

$$\hat{y}_{\max} = \frac{1}{h}\sum_{i=1}^{h} y_{\max}^{(i)} = \frac{1}{20}\left(y_{\min}^{(1)} + y_{\min}^{(2)} + \ldots + y_{\min}^{(20)}\right) = 227.7518512\ °C \quad (20)$$

$$s_{\hat{y}_{\max}} = \sqrt{\frac{1}{h(h-1)}\sum_{i=1}^{h=20}\left(y_{\max}^{(i)} - \hat{y}_{\max}\right)^2} = 2.48407 \times 10^{-6}\ °C \quad (21)$$

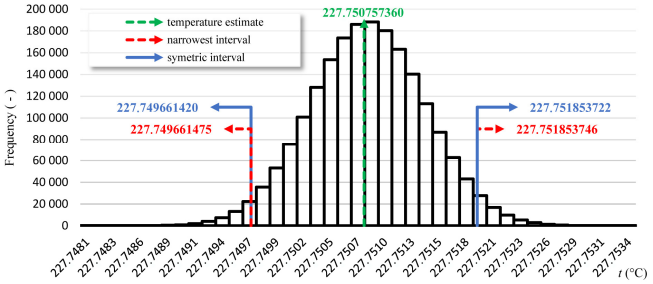

Fig.8. Histogram of output quantity *t*.

Numerical tolerance δ (see more in 5.C.), expressed by the standard uncertainty $\hat{u}(y)$, is

$$\delta = \frac{1}{2}\times 10^{-5} = 0.000005 = 5\times 10^{-6} \quad (22)$$

Stabilization criterion determines whether it is necessary to increase the current value of the sequence h>2 calculation of Monte Carlo method to the next sequence. This is the case if one of values $2s_{\hat{y}}$, $2s_{\hat{u}(y)}$, $2s_{\hat{y}_{\min}}$ and $2s_{\hat{y}_{\max}}$ becomes greater than δ. If the stabilization criteria are successfully met, $y^{(h\times M)}$, $u(y^{(h\times M)})$ are final statistical characteristics and $\left[y_{\text{low}}^{(h\times M)}\,; y_{\text{high}}^{(h\times M)}\right]$ are determined from all generated values. In the given example, following results were obtained by the Monte Carlo method:

Estimation of the temperature $t = 227.75075736$ °C, standard uncertainty $u(t) = 5.59884\times 10^{-4}$ °C, 95 % narrowest interval [227.74966148 ; 227.75185375] and 95 % symmetric interval [227.74966142 ; 227.75185372]. Statistical characteristic of output variable t, obtained by the adaptive Monte Carlo method, is shown in Fig.8. as a graph of probability distribution, where the model with one output variable t enters the submodel with multiple output variables a, b, c.

C. Validation of law of propagation of uncertainty

Calculation should prove if the reference interval obtained by law of propagation of uncertainty and by Monte Carlo method is identical in certain numerical tolerance. This numerical tolerance is assessed in relation to end points of the reference interval and it gives expression of standard uncertainty $u(y)$ to the existing number of decimal places. Numerical expression of tolerance δ with an associated standard uncertainty $u(y)$, as described in section 7.9.2 [4], is $\delta = \frac{1}{2}\times 10^{r}$, $u(y) = 56 \cdot 10^{-5}$ °C , $a = 56$, $r = -5$,

$$\Rightarrow \quad \delta = \frac{1}{2} \cdot 10^{-5} = 0{,}00005\ ^{\circ}\mathrm{C}$$

Absolute differences of limit values of both confidence intervals are determined as

$$d_{\mathrm{low}} = \left| y_{\mathrm{GUM}} - U_{0{,}95\ (\mathrm{GUM})} - y_{\mathrm{low(MCM)}} \right| \tag{23}$$

$$d_{\mathrm{high}} = \left| y_{\mathrm{GUM}} + U_{0{,}95\ (\mathrm{GUM})} - y_{\mathrm{high(MCM)}} \right| \tag{24}$$

Table 4. shows detailed result of uncertainty of temperature measurement according to ITS90 for each case listed. The maximum number of sequences needed for validation of each case was limited to 250. Table 3. contains validation results and minimum necessary number of sequences for all cases.

Fig.9. shows the user interface of created application for the evaluation and validation of the standard platinum resistance thermometer by the Monte Carlo method.

Table 3. Min. number of sequences h (each from $M = 10^5$) and validation result (yes ✓, no ×).

n	Ri (Ω)	Case a)	b)	c)	d)
1	24.8001933	2✓	2✓	2✓	2✓
2	25.7066394		2✓		2✓
3	26.6130855		5✓		
4	27.5195315		6✓		
5	28.4259776		12✓		
6	29.3324237		21✓		
7	30.2388697		27✓		
8	31.1453158		19✓		
9	32.0517618		20✓		
10	32.9582079		9✓		
11	33.8646540		36✓		
12	34.771100		35✓		250×
13	35.6775461		28✓		
14	36.5839922		31✓		
15	37.4904382		43✓		
16	38.3968843		24✓		
17	39.3033304	250×	38✓	250×	
18	40.2097764		41✓		
19	41.1162225		17✓		
20	42.0226686		23✓		
21	42.9291146		52✓		
22	43.8355607		20✓		9✓
23	44.7420067		27✓		16✓
24	45.6484528		16✓		17✓
25	46.5548989		20✓		34✓
26	47.4613449		28✓		34✓
27	48.3677910		25✓		34✓
28	49.2742371		41✓		34✓
29	50.1806831		26✓		51✓
30	51.0871292		26✓		
31	51.9935753		40✓		250×
32	52.9000213		32✓		
33	53.8064674		31✓		

n	Ri (Ω)	Case a)	b)	c)	d)
34	54.7129134		41✓		
35	55.6193595		45✓		
36	56.5258056		30✓		
37	57.4322516		20✓		
38	58.3386977		54✓		
39	59.2451438		43✓		
40	60.1515898		50✓		
41	61.0580359		106✓		
42	61.9644820		66✓		
43	62.8709280		45✓		
44	63.7773741		56✓		
45	64.6838202		59✓		
46	65.5902662		73✓		
47	66.4967123		46✓		
48	67.4031583		117✓		
49	68.3096044		58✓		250×
50	69.2160505	250×	76✓	250×	
51	70.1224965		85✓		
52	71.0289426		101✓		
53	71.9353887		79✓		
54	72.8418347		66✓		
55	73.7482808		82✓		
56	74.6547269		85✓		
57	75.5611729		61✓		
58	76.467619		250×		
59	77.3740651		89✓		
60	78.2805111		82✓		
61	79.1869572		128✓		
62	80.0934032		89✓		
63	80.9998493		151✓		
64	81.9062954		2✓		
65	82.8127414		2✓		124✓
66	83.7191875		2✓		124✓

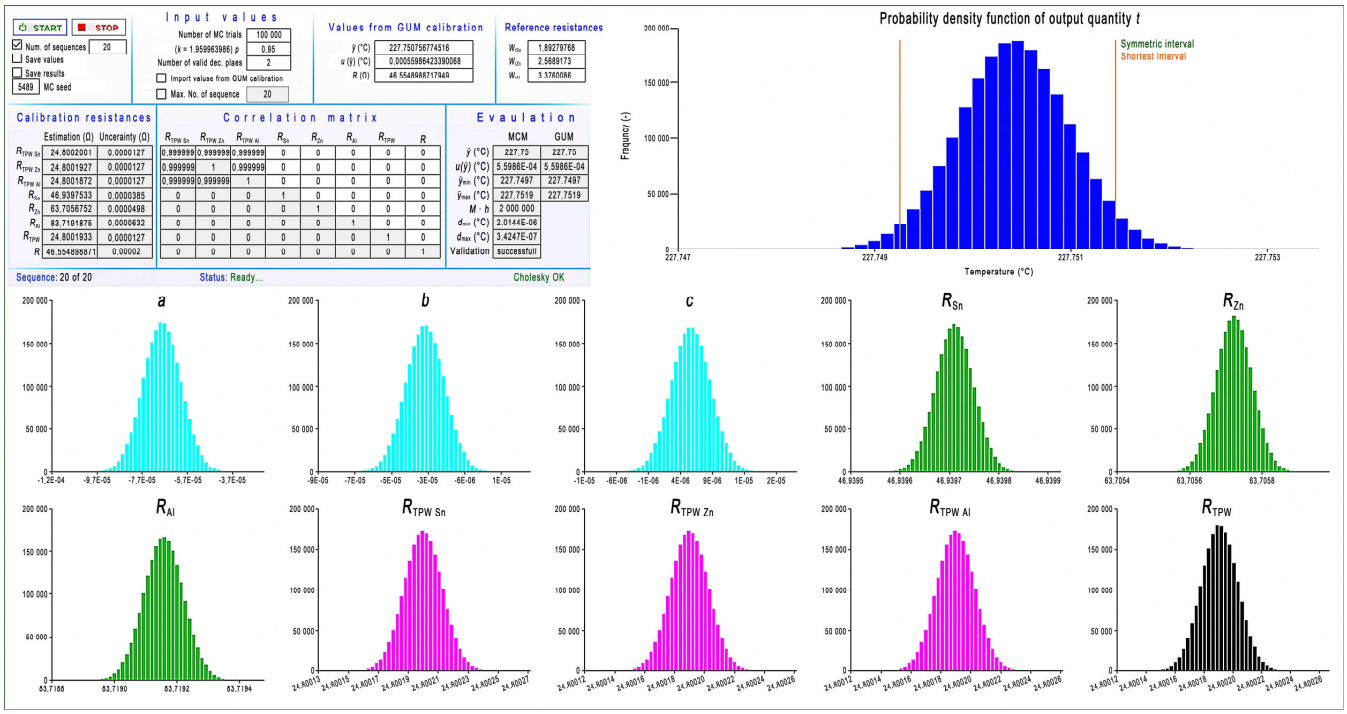

Fig.9. User interface of the application at the end of calculation.

Table 4. Comparison and verification of the law of uncertainty propagation through the law of distribution propagation using MCM for one value t in different cases (✓ = validation successful, ✖ = validation unsuccessful).

Case	Method	M ($\times 10^5$)	t (°C)	$u(t)$ (°C×10⁻⁴)	95% coverage interval Low / High	GUF validated
a)	GUF	-	227.750757	4.24	227.749925 / 227.751588	-
	MCM shortest	1	227.750758	4.98	227.749769 / 227.751722	✖
	MCM shortest	250	227.750757	5.60	227.749657 / 227.751852	✖
	MCM symmetric				227.749660 / 227.751855	✖
b)	GUF	-	227.750757	5.60	227.749659 / 227.751854	-
	MCM shortest	1	227.750758	5.60	227.749672 / 227.751863	✖
	MCM shortest	20	227.750757	5.60	227.749661 / 227.751854	✓
	MCM symmetric				227.749661 / 227.751854	✓
c)	GUF	-	227.750757	4.21	227.749931 / 227.751583	-
	MCM shortest	1	227.750758	4.76	227.749823 / 227.751687	✖
	MCM shortest	250	227.750757	4.75	227.749831 / 227.751692	✖
	MCM symmetric				227.749827 / 227.751688	✖
d)	GUF	-	227.750757	4.75	227.749826 / 227.751687	-
	MCM shortest	1	227.750758	4.76	227.749823 / 227.751687	✓
	MCM shortest	34	227.750757	4.75	227.749829 / 227.751690	✓
	MCM symmetric				227.749827 / 227.751688	✓
e)	GUF	-	227.750757	4.22	227.749929 / 227.751585	-
	MCM shortest	1	227.750758	4.98	227.749769 / 227.751722	✖
	MCM shortest	250	227.750757	4.98	227.749781 / 227.751732	✖
	MCM symmetric				227.749782 / 227.751733	✖
f)	GUF	-	227.750757	4.98	227.749781 / 227.751732	-
	MCM shortest	1	227.750758	4.98	227.749769 / 227.751722	✖
	MCM shortest	82	227.750757	4.98	227.749778 / 227.751729	✓
	MCM symmetric				227.749782 / 227.751733	✓

GUM uncertainty framework (GUF) [1], each sequence of MCM consist of M = 1×10^5 trials

6. Conclusions

This paper presents a procedure employing the Monte Carlo method for the determination of uncertainties of temperature scale. The procedure is based on generating pseudo-random numbers for the input SPRT resistances at DFPs and at TPW.

In order to consider the correlations among DFPs, the approach of generating pseudo-random numbers from multivariate distributions was used. To do so, an 8-dimensional Gaussian probability distribution was assumed. The assumption of Gaussian distribution is quite acceptable, given several sources of uncertainty of SPRT resistances at DFPs. If the correlations between the SPRT resistances at DFPs are negligible, it is possible to adapt the model so that the input resistances are uncorrelated and one-dimensional distributions for each input resistance can be used.

Fig.10. shows the course of uncertainty for the entire range considered and for each case. The figure compares uncertainties of temperatures obtained by using MCM and by law of propagation of uncertainty. As already mentioned in [16], the uncertainty due to the correlation between resistances of SPRT at DFPs can reduce the value of the uncertainty of temperature, doing so even in DFPs themselves.

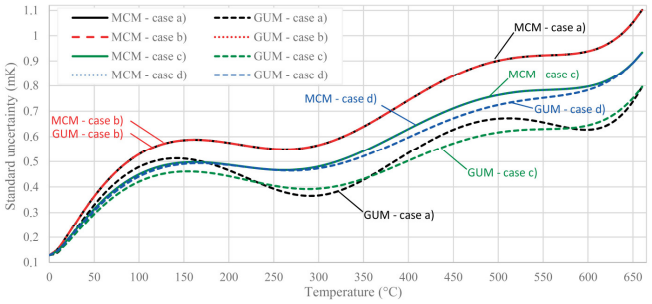

Fig.10. Comparison of MCM and GUM for different cases, illustrated on subrange (0 ÷ 660) °C of ITS-90.

Attention was also paid to validation of the use of law of propagation of uncertainty in accordance with the GUM for particular conditions. On the basis of validation, we found out that for some cases the results given by using MCM and law of propagation of uncertainty were not consistent.

Acknowledgment

Authors would like to thank the Slovak University of Technology in Bratislava, the APVV grant agency, projects number APVV15-0295, APVV 15-0164; the VEGA grant agency, projects number 1/0604/15, 1/0748/15; and the KEGA agency, projects number 014STU-4/2015 and 039STU-4/2017 for their support.

References

[1] Joint Committee for Guides in Metrology. (2008). *Evaluation of measurement data – Guide to the expression of uncertainty in measurement*. JCGM 100:2008.

[2] White, D.R., Ballico, M., Chimenti, V., et al. (2009). *CCT Working Document CCT/08-19/rev*.

[3] Joint Committee for Guides in Metrology. (2008). *Evaluation of measurement data – Supplement 1: "Guide to the expression of uncertainty in measurement" – Propagation of distributions using a Monte Carlo method*. JCGM 101:2008.

[4] Joint Committee for Guides in Metrology. (2011). *Evaluation of measurement data – Supplement 2: "Guide to the expression of uncertainty in measurement" – Extension to any number of output quantities*. JCGM 102:2011.

[5] Rosenkranz, P. (2011). Uncertainty propagation for platinum resistance thermometers calibrated according to ITS-90. *International Journal of Thermophysics*, 32 (1), 106-119.

[6] Duris, S., Palencar, R., Ranostaj, J. (2008). The efect of covariance on uncertainty when constructing the ITS-90 temperature scale. *Measurement Techniques*, 51 (4), 412-420.

[7] Duris, S., Palencar, R. (2006). A matrix interpretation of the estimate of the extension of uncertainties when constructing a temperature scale. *Measurement Techniques*, 49 (7), 689-696.

[8] Witkovsky, V. (2013). On exactmultiple-use linear calibration confidence intervals. In *MEASUREMENT 2013 : 9th International Conference on Measurement.* Bratislava, Slovakia : Institute of Measurement Science SAS, 35-38.

[9] Lira, I., Grientschnig, D. (2013). A formalism for expressing the probability density functions of interrelated quantities. *Measurement Science Review*, 13 (2), 50-55.

[10] Kosarevsky, S.V., Latypov, V.N. (2012). Practical procedure for position tolerance uncertainty determination via Monte-Carlo error propagation. *Measurement Science Review*, 12 (1), 1-7.

[11] Zhang, F. Qu, X. (2012). Fusion estimation of point sets from multiple stations of spherical coordinate instruments utilizing uncertainty estimation based on Monte Carlo. *Measurement Science Review*, 12 (2), 40-45.

[12] Ribeiro, S., Alves, J., Oliveira, C., Pimenta, M., Cox, M.G. (2008). Uncertainty evaluation and validation of a comparison methodology to perform in-house calibration of platinum resistance thermometers using a Monte Carlo method. *International Journal of Thermophysics*, 29 (3), 902-914.

[13] Palenčár, R., Ďuriš, S., Dovica, M., Palenčár, J. (2015). Application of Monte Carlo method for evaluation of uncertainties of temperature measurement by SPRT. In *XXI IMEKO World Congress "Measurement in Research and Industry"*. IMEKO, 6 p.

[14] Lira, I. Grientschnig, D. (2015). Bayesian analysis of a simple measurement model distinguishing between types of information. *Measurement Science Review*, 15 (6), 274-283.

[15] Preston-Thomas, H. (1990). The International temperature scale of 1990 (ITS-90). *Metrologia*, 27 (1), 3-10.

[16] Palencar, R., Duris, S., Durisova, Z., Brokes, V., Pavlasek, P. (2016). Reduction of measurement uncertainty by taking into account correlation in measurements and temperature scale realization. In *Measurement Techniques*, 59 (1), 52-58.

Received December 12, 2016.
Accepted May 09, 2017.

MEASUREMENT SCIENCE REVIEW, **17**, (2017), No. 3, 117-124

DE GRUYTER OPEN

MEASUREMENT SCIENCE REVIEW

Journal homepage: http://www.degruyter.com/view/j/msr

ISSN 1335-8871

Stationary Wavelet-based Two-directional Two-dimensional Principal Component Analysis for EMG Signal Classification

Yi Ji[1], Shanlin Sun[2], Hong-Bo Xie[3]

[1]*School of Electrical & Information Engineering, Jiangsu University, Zhenjiang 212013, China*
[2]*School of Instrumentation Science & Opto-Electronics Engineering, Beihang University, Beijing 100191, China*
[3]*ARC Centre of Excellence for Mathematical & Statistical Frontiers, Queensland University of Technology, Brisbane 4001, Australia, hongbo.xie@qut.edu.au*

Discrete wavelet transform (WT) followed by principal component analysis (PCA) has been a powerful approach for the analysis of biomedical signals. Wavelet coefficients at various scales and channels were usually transformed into a one-dimensional array, causing issues such as the curse of dimensionality dilemma and small sample size problem. In addition, lack of time-shift invariance of WT coefficients can be modeled as noise and degrades the classifier performance. In this study, we present a stationary wavelet-based two-directional two-dimensional principal component analysis (SW2D^{2}PCA) method for the efficient and effective extraction of essential feature information from signals. Time-invariant multi-scale matrices are constructed in the first step. The two-directional two-dimensional principal component analysis then operates on the multi-scale matrices to reduce the dimension, rather than vectors in conventional PCA. Results are presented from an experiment to classify eight hand motions using 4-channel electromyographic (EMG) signals recorded in healthy subjects and amputees, which illustrates the efficiency and effectiveness of the proposed method for biomedical signal analysis.

Keywords: Wavelet transform, principal component analysis, feature extraction, pattern classification, electromyographic signal.

1. INTRODUCTION

Biomedical signal analysis has been broadly applied for robotics control, human/brain machine interface, disease diagnosis, wearable devices, and rehabilitation programming. Most biomedical signals, for example electromyography (EMG), an electrical manifestation of skeletal muscle contractions, are typically nonlinear and nonstationary. Discrete wavelet transform (WT) offers simultaneous interpretation of the EMG signal in both time and frequency domains which allows to elucidate local, transient or intermittent components at various scales [1]. However, there are typically a large amount of wavelet coefficients generated from such two-dimensional time-frequency (TF) analysis. In addition to noise interferences, irrelevant or redundant information may exist in the wavelet coefficients. Principal component analysis (PCA) decomposes the covariance structure of the dependent variables into orthogonal components by calculating the eigenvalues and eigenvectors of the data covariance matrix. It linearly projects the original data in a high-dimensional space to a set of uncorrelated components in a low-dimensional feature space while preserving the most original information at the same time. Therefore, WT combined with PCA (WTPCA) has been one of the most powerful approaches to simultaneously extract discriminative feature and reduce the dimension in the EMG study. The basic routine of this hybrid method consists of decomposition of EMG signals into time-frequency plane, rearrangement of the time-frequency elements into a row vector, and reduction of the dimension using PCA. Englehart et al. [2] decomposed four channel transient EMG signals by short-time Fourier transform (STFT), WT, and wavelet packet transform (WPT) into TF plane to discriminate six hand motions for prosthetic hand control. They compared the performance of PCA feature reduction and Euclidean distance class separability (CS) criterion. The results indicated that PCA was vastly superior to CS dimensionality reduction, as well as significantly improving the WT and WPT-based methods in comparison with time domain feature when using linear discriminant analysis classifier. Khezri and Jahed's study using adaptive neuro-fuzzy inference system further confirmed the superiority of WT-PCA hybridization in EMG-based hand motion pattern recognition [3]. Qi et al. [4], [5] utilized the principal components of EMG intensity spectra obtained from nonlinearly-scaled wavelets to compare motor unit recruitment patterns during isometric ramp and step muscle contractions, as well as dynamic concentric and eccentric contractions of the human biceps brachii. The same WT-PCA scheme was also employed to discriminate between

DOI: 10.1515/msr-2017-0015

fast and slow muscle fibers [6], to investigate motor unit recruitment patterns between and within muscles of the dysfunction in children and young adults with cerebral palsy [8]. Weiderpass et al. [9] investigated the alternations of thigh and calf muscles recruitment strategies during gait among non-diabetic and diabetic neuropathic patients by using an adaptive optimal kernel time-frequency representation and discrete WT followed by PCA. In all of these WTPCA-based EMG representation and recognition methods, WT coefficients at various scales must be first transformed into a vector. However, concatenating WT coefficients at various scales into a 1D array often leads to a high-dimensional vector space, where it is difficult to evaluate the covariance matrix accurately due to its large size and the relatively small number of training samples. Furthermore, computing the eigenvectors of a large size covariance matrix is very time-consuming, whilst the response time of EMG real-time control systems should not introduce a delay that is perceivable by the user [1].

In fact, a two-dimensional WT coefficient matrix can be regarded as an image. It is thus feasible to apply image processing techniques to indicate the WT coefficient matrix characteristics. Two-dimensional principal component analysis (2DPCA) developed by Yang et al. [10] is a 2D image representation and reduction technique, in which an image matrix does not need to be transformed into a 1D array. Many experimental results have indicated that 2DPCA is computationally more efficient than PCA in the extraction of image features. Although 2DPCA is typically able to obtain higher recognition accuracy than PCA, a vital unresolved problem is that 2DPCA needs many more coefficients for image or TF matrix representation than PCA [11]. Zhang and Zhou [11] indicated that 2DPCA essentially operates along the row direction of the image matrix and, thus, proposed an alternative 2DPCA operating along the column direction. By simultaneously considering the row and column directions, they developed the two-directional triceps muscles [7], as well as quantify dynamic muscle two-dimensional principal component analysis ($2D^2PCA$) for a more efficient image representation and recognition.

Another issue of discrete wavelet transform is the lack of time-shift invariance caused by down sampling by two. Since all even-indexed outputs of a half-band filter are discarded, a small shift of the input signal causes a large change in the WT sub-band coefficients. Lack of time-shift invariance of WT coefficients can be modeled as noise and degrades the classifier performance [12]. The stationary wavelet transform (SWT) does not decimate the signal at each stage, as does the standard discrete WT, avoiding the problem of nonlinear distortion of the WT with shifts in the signal.

Inspired by the success of $2D^2PCA$ in imaging processing and time-invariant characteristics of SWT, the purpose of this study is to develop an efficient and effective feature extraction method for fully exploiting the time-frequency information of biomedical signals. The size of the SWT covariance matrix is equal to the length/width of time-frequency plane in $2D^2PCA$, which is quite smaller than the size of a covariance matrix in PCA. The evaluation of covariance matrix is thus more accurate and the estimation of corresponding eigenvectors is more efficient than PCA. The key idea is that $2D^2PCA$ is applied to reduce the dimension of SWT coefficient matrix in a highly efficient manner for pattern classification. The method is, therefore, termed as stationary wavelet-based two-directional two-dimensional principal component analysis ($SW2D^2PCA$). To illustrate the efficiency and effectiveness of the proposed method, results are presented on the recognition of eight hand motions from 4-channel EMG signals recorded in both healthy subjects and amputees.

2. SUBJECT & METHODS

2.1. Stationary wavelet transform

The wavelet transform of a function *f*, with respect to a given mother wavelet ψ , is defined as

$$w_s f(x) = f * \psi_s(x) = \frac{1}{s}\int_{-\infty}^{\infty} f(t)\psi\left(\frac{x-t}{s}\right)dt, \tag{1}$$

where *s* is the scale factor. Assume $s = 2^j$ ($j \in R$, *R* is the integral set), the dyadic WT can be represented as

$$S_{2^j} f(n) = \sum_{k\in R} h_k S_{2^{j-1}} f(n-2^{j-1}k), \tag{2}$$

$$W_{2^j} f(n) = \sum_{k\in R} g_k S_{2^{j-1}} f(n-2^{j-1}k), \tag{3}$$

where S_{2^j} is a smoothing operator and $W_{2^j} f(n)$ is the WT of the discrete signal *f*(*n*), whilst h_k and g_k are the coefficients of corresponding low-pass and high-pass filters, respectively. The standard discrete wavelet transform decimates the wavelet coefficients at each scale, resulting in the half size of the original series. On the other hand, the stationary wavelet transform pads the corresponding low-pass and high-pass filters with zeros between coefficients at each scale. Two new sequences at sub-band thus have the same size as the original sequence. The major advantage of SWT is the preservation of time information of the original signal sequence at each level, particularly useful for feature extraction and denoising [1].

2.2. $2D^2PCA$ schematic diagram

Fig.1. is a schematic diagram of $2D^2PCA$. Without loss of generality, we consider an *m* by *n* time-frequency matrix (TFM) $\mathbf{A}$ obtained from the stationary wavelet decomposition. Let $\mathbf{X} \in ¡^{n\times q}$ and $\mathbf{Y} \in ¡^{m\times p}$ be matrices having orthonormal columns $n \times q$ and $m \times p$, respectively. We can simultaneously project $\mathbf{A}$ onto $\mathbf{X}$ to yield the $m \times q$ matrix $\mathbf{B} = \mathbf{AX}$, and onto $\mathbf{Y}$ to yield the $p \times n$ matrix $\mathbf{C} = \mathbf{Y}^T\mathbf{A}$. In contrast to conventional PCA for one-dimensional array applications, $2D^2PCA$ operates on a matrix in both horizontal and vertical directions. The total scatter of the projected samples, a measure of the discriminatory power of a projection matrix, can be characterized by its trace of the covariance matrix of the projected matrix. From this point of

view, maximization of the generalized total scatter is the criterion adopted to find the optimal projection matrices $\mathbf{X}$ and $\mathbf{Y}$ for row and column directions, respectively:

$$\begin{aligned} J(\mathbf{X}) &= tr\{E[(\mathbf{B}-E(\mathbf{B}))(\mathbf{B}-E(\mathbf{B}))^T]\} \\ &= tr\{E[(\mathbf{AX}-E(\mathbf{AX}))(\mathbf{AX}-E(\mathbf{AX}))^T]\} \\ &= tr\{\mathbf{X}^T E[(\mathbf{A}-E(\mathbf{A}))^T(\mathbf{A}-E(\mathbf{A})]\mathbf{X}\}, \end{aligned} \tag{4}$$

$$\begin{aligned} J(\mathbf{Y}) &= tr\{E[(\mathbf{C}-E(\mathbf{C}))(\mathbf{C}-E(\mathbf{C}))^T]\} \\ &= tr\{E[(\mathbf{Y}^T\mathbf{A}-E(\mathbf{Y}^T\mathbf{A}))(\mathbf{Y}^T\mathbf{A}-E(\mathbf{Y}^T\mathbf{A}))^T]\} \\ &= tr\{\mathbf{Y}^T E[(\mathbf{A}-E(\mathbf{A}))(\mathbf{A}-E(\mathbf{A}))^T]\mathbf{Y}\}, \end{aligned} \tag{5}$$

where $tr\{\bullet\}$ is the trace.

Considering the $m \times q$ matrix $\mathbf{B} = \mathbf{AX}$ obtained by projecting $\mathbf{A}$ onto $\mathbf{X}$ in (4), the horizontal covariance matrix is denoted by

$$\mathbf{G}_h = E[(\mathbf{A}-E(\mathbf{A}))^T(\mathbf{A}-E(\mathbf{A}))], \tag{6}$$

which is an $n \times n$ positive semi-definite matrix.

Suppose that the training feature set is $\mathbf{\Omega} = (\mathbf{A}_1, \mathbf{A}_2, \mathrm{L}\ , \mathbf{A}_N)$, where each $\mathbf{A}_i (i = 1,2,\mathrm{L}\ ,\mathbf{N})$ denotes the ith $m \times n$ time-frequency matrix and $\mathbf{N}$ is the number of training samples. The average TFM is given by

$$\bar{\mathbf{A}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{A}_i \tag{7}$$

Denoting the kth row vectors of $\mathbf{A}_i$ and $\bar{\mathbf{A}}$ by $\mathbf{A}_i^k$ and $\bar{\mathbf{A}}_h^k$, respectively, these TFMs can be represented by

$$\mathbf{A}_i = [(\mathbf{A}_i^1)^T, (\mathbf{A}_i^2)^T, \mathrm{L}\ , (\mathbf{A}_i^m)^T]^T, \tag{8}$$

and

$$\bar{\mathbf{A}} = [(\bar{\mathbf{A}}_h^1)^T, (\bar{\mathbf{A}}_h^2)^T, \mathrm{L}\ , (\bar{\mathbf{A}}_h^m)^T]^T. \tag{9}$$

The horizontal covariance matrix can then be obtained from the outer product of these TFM row vectors:

$$\begin{aligned} \mathbf{G}_h &= \frac{1}{N}\sum_{i=1}^{N}(\mathbf{A}_i-\bar{\mathbf{A}})^T(\mathbf{A}_i-\bar{\mathbf{A}}) \\ &= \frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{m}(\mathbf{A}_i^k-\bar{\mathbf{A}}_h^k)^T(\mathbf{A}_i^k-\bar{\mathbf{A}}_h^k) \end{aligned} \tag{10}$$

Similarly, for the $p \times n$ matrix $\mathbf{C} = \mathbf{Y}^T\mathbf{A}$ obtained by projecting $\mathbf{A}$ onto $\mathbf{Y}$ in (5), the vertical covariance matrix can be denoted by

$$\mathbf{G}_v = E[(\mathbf{A}-E(\mathbf{A}))(\mathbf{A}-E(\mathbf{A}))^T], \tag{11}$$

which is $m \times m$ positive semi-definite matrix.

TMFs and their average are now denoted by column vectors:

$$\mathbf{A}_i = [(\mathbf{A}_i^1)^T, (\mathbf{A}_i^2)^T, \mathrm{L}\ , (\mathbf{A}_i^n)], \tag{12}$$

$$\bar{\mathbf{A}} = [(\bar{\mathbf{A}}_v^1)^T, (\bar{\mathbf{A}}_v^2)^T, \mathrm{L}\ , (\bar{\mathbf{A}}_v^n)]. \tag{13}$$

where $\mathbf{A}_i^j$ and $\bar{\mathbf{A}}_v^j$ denote the jth column vectors of $\mathbf{A}_i$ and $\bar{\mathbf{A}}$, respectively.

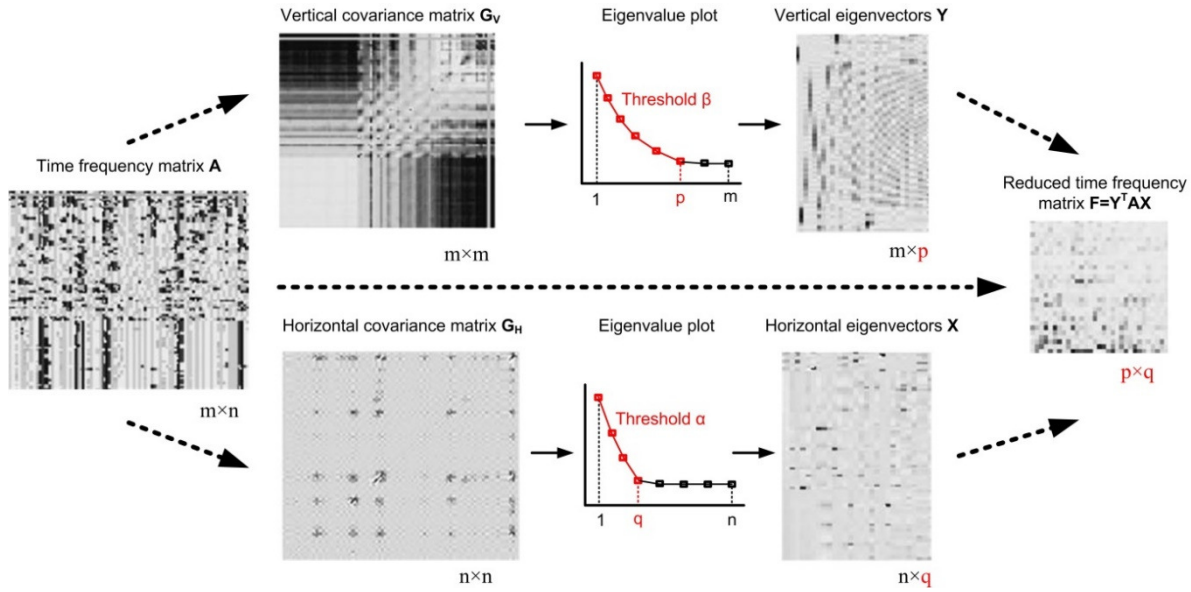

Fig.1. Schematic diagram of two-directional two-dimensional principal component analysis to obtain the reduced time-frequency matrix F (right) from an input time-frequency matrix A (left).

Now, the vertical covariance matrix of (11) can be constructed from the outer products of column vectors:

$$\begin{aligned}\mathbf{G}_v &= \frac{1}{N}\sum_{i=1}^{N}(\mathbf{A}_i - \bar{\mathbf{A}}_h)(\mathbf{A}_i - \bar{\mathbf{A}}_h)^T \\ &= \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{n}(\mathbf{A}_i^j - \bar{\mathbf{A}}_v^j)(\mathbf{A}_i^j - \bar{\mathbf{A}}_v^j)^T\end{aligned} \tag{14}$$

Zhang and Zhou [11] demonstrated that the optimal projection matrices $\mathbf{X}$ and $\mathbf{Y}$ are composed of the orthonormal eigenvectors $\mathbf{X}_1, \mathbf{X}_2, \mathrm{L}\ , \mathbf{X}_q$ of $\mathbf{G}_h$ corresponding to the q largest eigenvalues and $\mathbf{Y}_1, \mathbf{Y}_2, \mathrm{L}\ , \mathbf{Y}_p$ of $\mathbf{G}_v$ corresponding to the p largest eigenvalues, respectively. The values of p and q can be controlled by two pre-set thresholds, and, corresponding to the energy conservation rates at two directions. In practice, we set a ratio of total energy preserved as in PCA (for example >85 %) and then set $\alpha = \beta$ [13], [14].

After obtaining the projection matrices $\mathbf{X}$ and $\mathbf{Y}$, 2D^{2}PCA projects the m by n TFM $\mathbf{A}$ onto $\mathbf{X}$ and $\mathbf{Y}$ simultaneously, yielding the reduced p by q matrix

$$\mathbf{F} = \mathbf{Y}^T\mathbf{A}\mathbf{X}. \tag{15}$$

Using the above procedure, an $m \times n$ dimensional feature matrix $\mathbf{A}$ is projected into a $p \times q$ dimensional feature matrix $\mathbf{F}$.

2.3. SW2D^{2}PCA schematic diagram

In this section, we describe the stationary wavelet-based two directional two-dimensional principal component analysis algorithm for extracting discriminant feature information from these matrices as follows:

1. Multiple-channel signals are first segmented by a moving window with width d. Choose a time-frequency decomposition method. That is, specify the mother wavelet function and decomposition level. The stationary wavelet transform is then employed to decompose each time-segment of individual channels into details $D_1, D_2, \mathrm{L}\ , D_L$ and approximate A_L under the same decomposition level L.
2. 2D^{2}PCA is subsequently carried out on each of the $d \times (L+1)$ dimension matrices to extract the most informative features, as well as reduce the dimension based on the user-specified threshold of total energy preserved.
3. Since the discriminant abilities of principal components (PCs) at various scales are different, a simple distance-based technique is applied to re-order all PCs [1].
4. The performance of the algorithm is evaluated by feeding the optimal PCs obtained into a classifier.

2.4. Experimental protocol and performance evaluation

The proposed algorithm was evaluated using the EMG data collected from the following experiment. Eight distinct wrist and hand motions were used: grasp (GR), hand open (OP), wrist flexion (WF), wrist extension (WE), ulnar deviation (UD), radial deviation (RD), pinch (PN), and thumb flexion (TF), as depicted in Fig.2. These represent the commonly used wrist and hand movements in daily life.

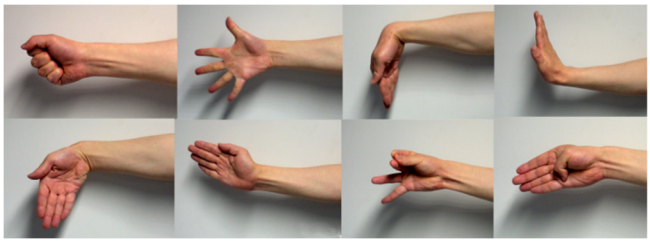

Fig.2. Eight classes of motion were used in the experiment. From the left to right in the first row: grasp (GP), hand open (OP), wrist flexion (WF), wrist extension (WE), and in the second row: ulnar deviation (UD), radial deviation (RD), pinch (PN), thumb flexion (TF).

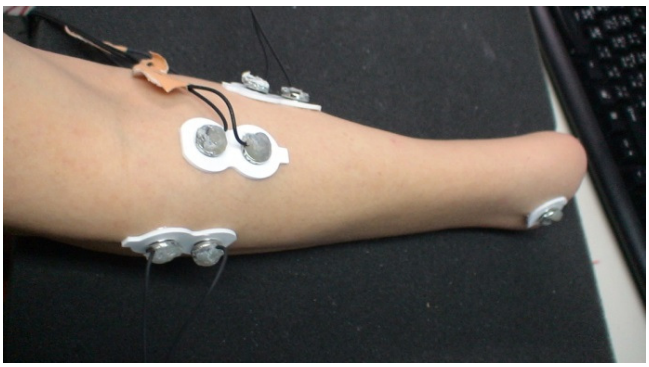

Fig.3. The experimental setting to record EMG signals from an amputee. There are four pairs of electrodes on the forearm with one pair under the forearm. The electrode at the wrist provides the common ground reference.

In the experiment, the EMG data were collected from ten healthy subjects and two amputees (eight males and four females, 30 ± 6.8 years). The human subject ethical approval was obtained from the relevant committee and informed consent was obtained from all subjects prior to the experiment. Four channels of EMG signals were acquired from the forearm using the EMG bi-polar Ag-AgCl electrodes (Dual electrode #272, Noraxon USA Inc. AZ, USA). Electrodes were placed on the extensor digitorum, the extensor carpi radialis, the palmaris longus and the flexor carpi ulnaris around the forearm. The distance of two surface electrodes was 2 cm. Skin areas of interest were abrased beforehand with alcohol. An additional Ag-AgCl electrode was placed on the wrist to provide a common ground reference. Fig.3. is the experimental setting for an amputee

with three pairs of electrodes being visible and another pair invisible due to its placement on the other side of the forearm. EMG signal was amplified by an amplifier (RM-6280C, Chengdu Device Inc. Sichuan, China) with a gain of 2000, filtered by 8–500 Hz band-pass analog filter within the amplifier, and then digitized by a 12-bit data acquisition card (NI PCI-6024E, National Instruments, Austin, TX) with the sample frequency of 1 kHz.

Fifteen sessions were conducted for each subject. The first five sessions were used for the learning procedures, while the sixth to tenth session as the validation set and the remaining for performance evaluations. Each subject was asked to maintain a static contraction for each motion and to change the motions with a fixed movement velocity. For those specific tasks the amputees cannot perform, they tried to perform under the guidance. In every session, each motion was performed once for a duration of 5 s, then switched to another motion in random order.

The 4-channel EMG data were further segmented into a series of overlapping windows (window length: 256 ms, overlap step: 128 ms). Since the Symmlet-5 has been proven to be an effective mother wavelet for stationary wavelet-based classification [15], it was selected to simultaneously decompose EMG signals over six levels. The remaining procedures for SW2D^{2}PCA described in Section 2.3 were employed to extract two-dimensional PCs. Support vector machine (SVM), a typical nonlinear EMG classifier used in previous study [1]-[3], was employed to evaluate the classification performance of the proposed algorithm. After the classification, the accuracy was further improved by a post-processing procedure using majority vote (MV) [16]. Conventional WTPCA algorithm to analyse the same data set was also devised for comparison.

3. RESULTS

3.1. Multi-scale muscle activity patterns

Using the proposed SW2D^{2}PCA technique, the EMG signal at each channel was first transformed into a two-dimensional matrix. Fig.4. shows the typical contour plots for eight motions for subject 3, each row corresponding to a motion type. With each intended motion, a significant difference between the intensity of the surface EMG signals over the upper limb muscles can be readily discerned in the first column contour plots. Similar to the panels in the first column, there was significant discrepancy in the intensity distributions of the remaining contour plots in the remaining three columns, indicating useful discriminant information in the SWT matrices.

The two-directional two-dimensional principal component analysis was then used to reduce the dimension of each matrix. Fig.5. shows the contour plots of each matrix in Fig.4. following dimension reduction using 2D^{2}PCA when the energy conservation rate and total energy preserved were 98 % and 90 %, respectively. Compared with Fig.4., the intensity difference between certain sub-panels in Fig.5. is further enhanced, including, for example, those in the first row. In addition, the matrix size at each channel was significantly decreased, which were 80×3, 86×3, 122×4, and 116×4, respectively. If conventional PCA was used with all time-frequency coefficients arranged into a 1D array, the size of the covariance matrix would be 1792×1792. However, the use of 2D^{2}PCA resulted in the size of all covariance matrices being less than 130×130, avoiding the curse of dimensionality and small sample issue as well as improving the numerical stability. It should be noted that the reduced dimension of each channel for all subjects were different because EMG signal varied from subject to subject due to the physiological factors.

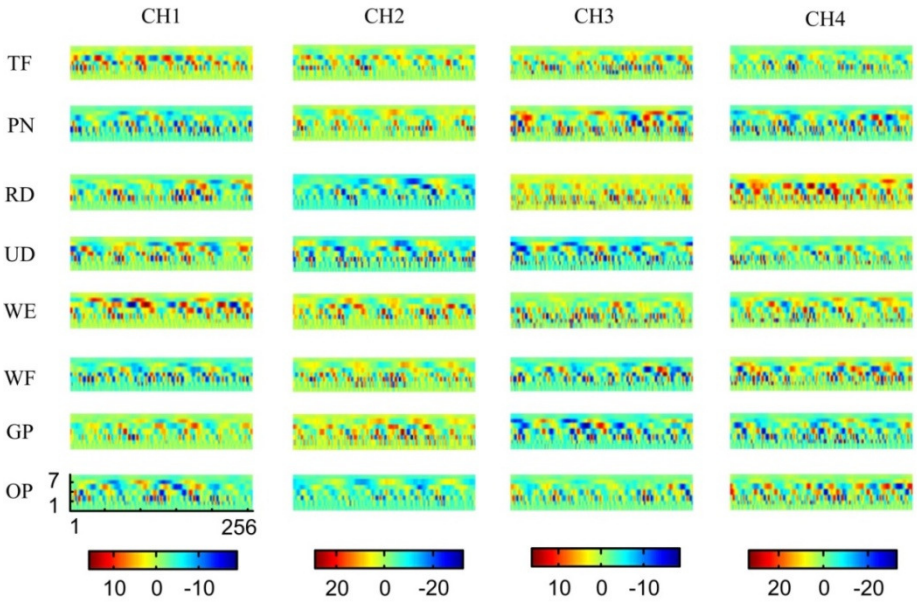

Fig.4. Contour plots of stationary wavelet transform matrices for 4-channel EMG traces of eight hand motions obtained from subject 3. The abscissa represents the time and the ordinate represents the frequency or the scale of stationary wavelet transform. The color bar indicates the strength of the muscle electrical activity.

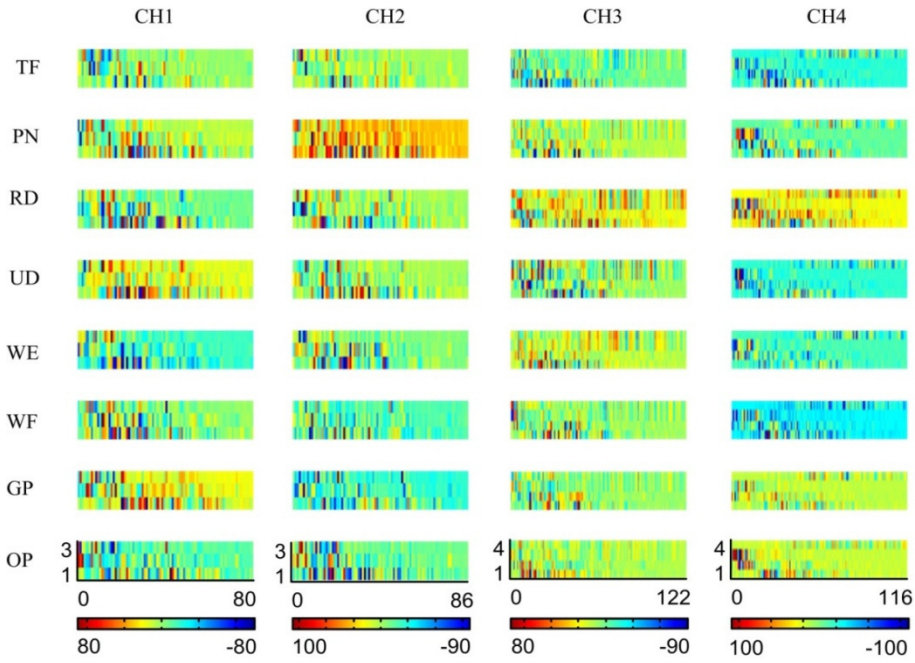

Fig.5. The contour plots of stationary wavelet transform matrices reduced using 2D^{2}PCA for 4-channel EMG signals of eight hand motions obtained from subject 3. The abscissa and ordinate represent the reduced size of the contour plots in Fig.4. The color bar indicates the relative strength of the muscle myoelectric activity after dimension reduction.

3.2. Effect of energy conservation rate

A large energy conservation rate results in more information loss, whilst a low rate increases the computational burden. To reach a trade-off between these two factors, three energy conservation rates of 97 %, 98 % and 99 %, were employed to assess its effect on classification accuracy. Fig.6. shows the average accuracy across the 12 subjects at these various energy conservation rates for the SVM classifier. With the increasing number of PCs, the accuracy of all three conservation rates initially increased and then entered a relatively flat range with moderate fluctuations. The optimal PCs to achieve the highest accuracy for three conservation rates were all in the range from 25 to 35. In addition, there was no significant difference between energy conservation rates (p<0.01).

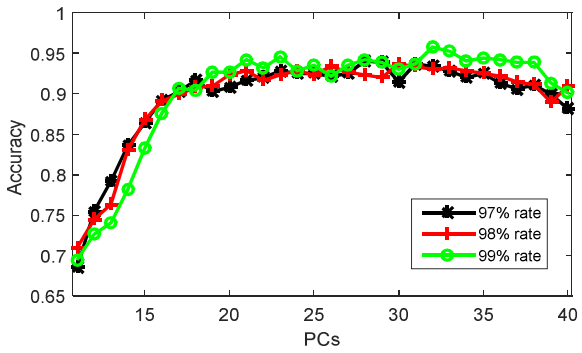

Fig.6. The effect of 97 % (black), 98 % (red), and 99 % (green) energy conservation rate of SW2D^{2}PCA on the EMG signals classification accuracy.

3.3. Effect of total energy conserved

For PCA analysis, a typical recommendation is to set the threshold of total energy conserved between 0.8 and 0.95. Fig.7. shows the classification accuracy for SVM for three threshold values of total energy conserved, i.e., 95 %, 90 %, and 85 %. With the reduction in threshold, there was no significant difference in the accuracy for SVM. The insensitivity of SVM to the total energy preserved may be due to its adaptive ability to map input features to high-dimensional feature space.

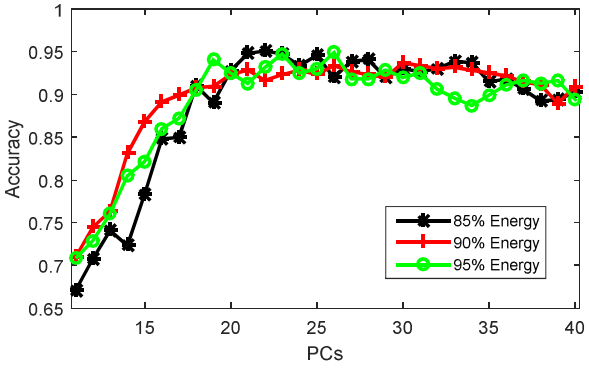

Fig.7. The effect of 85 % (black), 90 % (red), and 95 % (green) total energy conserved of SW2D^{2}PCA on the EMG signals classification accuracy.

3.4. Recognition of intended motions

Pattern recognition analysis was performed using the optimal number of PCs previously determined. Table 1. summarizes the subject-specific classification accuracy for

all eight intended upper-limb motions. An average classification accuracy above 93 % could be achieved among all subjects after majority vote. Across all subjects, there is significant difference between the accuracy of SW2D^{2}PCA and WTPCA (p<0.05) with lower average accuracy for WTPCA in both cases - with or without majority vote. On the other hand, because the amputees can only perform grasping, opening based on imagination, the related muscle activities were not as strong as in the healthy subjects. The accuracy for two amputees was much lower than for the healthy subjects. As mentioned before, it should be emphasized that EMG activity is subject-dependent for both healthy subjects and amputees. Therefore, the structure and information distribution in the time-frequency matrices varied between subjects, which led to different reduced sizes with SW2D^{2}PCA. Ultimately, this subject-specific time-frequency distribution of EMG feature information led to inconsistent classification errors among different subjects. The subject-specific EMG activity and classification performance suggested that optimal myoelectric pattern-recognition control system parameters should be individually customized.

Table 1. Classification results of all twelve subjects by proposed stationary wavelet two-directional two-dimensional principal component analysis and conventional wavelet principal component analysis based feature subsets

Subject	Before MV		After MV	
	SW2D^{2}PCA	WTPCA	SW2D^{2}PCA	WTPCA
1	92.15	83.52	98.28	89.31
2	88.92	85.19	95.66	94.28
3	89.50	90.54	95.33	95.95
4	91.65	86.09	96.64	92.73
5	89.98	80.94	94.73	84.06
6	91.11	84.53	97.86	91.69
7	93.39	88.75	98.23	96.35
8	91.91	83.28	94.31	87.92
9	92.27	91.04	99.95	96.00
10	89.93	82.58	95.57	87.76
11*	79.74	72.55	86.11	77.78
12*	72.29	68.16	78.47	71.09
Average	88.57±6.2	83.09±6.8	94.26±6.1	88.74±7.8

4. DISCUSSION / CONCLUSIONS

A novel stationary wavelet-based two-directional two-dimensional principal component analysis for signal classification has been proposed and examined in this study. One of the major challenges related to the design of EMG interfaces is to maintain high classification accuracy in long-term use [17]. In real use, the muscle contractions, i.e. the classes associated to control commands, are performed in a variety of conditions, which may lead to differences in signal properties making them unrecognizable for the classifier. Stationary wavelet transform avoids the problem of nonlinear distortion of the wavelet and wavelet package transforms with shifts in the signal. Recently, with the improvements in physiological measurement equipment for EMG as well as electroencephalography and magnetoencephalography, new technology permits registration of up to several hundred channels using high-density electrode arrays. Such arrays with small electrode sizes and inter-electrode spacing can cover large areas of the tissue, providing extra spatial information which is largely independent of any "classical" temporal information. To effectively and efficiently extract feature from such high-dimensional signal space is another challenge in biosignals analysis and their applications [18]. Compared with the existing PCA method, 2D^{2}PCA provides an improved approach to extract discriminative feature and reduce the dimension from the high-dimensional random and complex raw signals.

In order to test this approach, we used SW2D^{2}PCA to extract and classify specific TF patterns in four-channel EMG signals from ten healthy subjects and two amputees for identification of eight hand motions. SW2D^{2}PCA achieved higher accuracy than WTPCA for both healthy subjects and amputees before and after majority vote. For the healthy subjects, the average accuracy exceeds 96.6 %, which can be employed as a promising technique for human-machine interaction or robot control. However, in comparison to the healthy subjects, the classification accuracy for two amputees using SW2D^{2}PCA is relatively low although it is higher than that of WTPCA. This is due to the fact that the amputees can only perform grasping and opening based on imagination, the relevant muscle activities were not as strong as in the healthy subjects. In this study, Symmlet-5 was employed as the mother wavelet to decompose EMG signals for all subjects. Studies have indicated that signal matched or optimized wavelet can substantially enhance the classification accuracy [19], [20]. Another limitation of this study is that the PCs of each channel obtained using 2D^{2}PCA are re-ordered and further reduced using a simple distance measure, which is equivalent to a two-step reduction method. It is necessary to develop a unified framework to simultaneously extract discriminative features from multiple channels. In addition, effective training can improve the electrical activities of residual muscles, and, therefore decrease the recognition error for amputees. It is expected that integration of these measures will result in enhanced pattern recognition of motion patterns for the amputees. The efficiency and effectiveness of the method can be further validated by using high-dimensional EEG, MEG, and fMRI signals. Although the present study focuses on signal pattern classification, based on the PCs obtained from time-frequency plane, it is relatively straightforward to expand SW2D^{2}PCA for signal compression, denoising, instantaneous frequency estimation, and other related tasks.

REFERENCES

[1] Xie, H.B., Zheng, Y.P., Guo, J.Y. (2009). Classification of the mechanomyogram signal using a wavelet packet transform and singular value decomposition for multifunction prosthesis control. *Physiological Measurement,* 30, 441-457.

[2] Engelhart, K., Hudgins, B., Parker, P.A., Stevenson, M. (1999). Classification of the myoelectric signal using time-frequency based representations. *Medical Engineering & Physics*, 21, 431-438.

[3] Khezri, M., Jahed, M. (2007). Real-time intelligent pattern recognition algorithm for surface EMG signals. *Biomedical Engineering Online*, 6, 45.

[4] Qi, L.P., Wakeling, J.M., Ferguson-Pell, M. (2011). Spectral properties of electromyographic and mechanomyographic signals during dynamic concentric and eccentric contractions of the human biceps brachii muscle. *Journal of Electromyography and Kinesiology*, 21, 1056-1063.

[5] Qi, L.P., Wakeling, J.M., Green, A., Lambrecht, K., Ferguson-Pell, M. (2011). Spectral properties of electromyographic and mechanomyographic signals during isometric ramp and step contractions in biceps brachii. *Journal of Electromyography and Kinesiology*, 21, 128-135.

[6] Von Tscharner, V., Goepfert, B. (2006). Estimation of the interplay between groups of fast and slow muscle fibers of the tibialis anterior and gastrocnemius muscle while running. *Journal of Electromyography and Kinesiology*, 16, 188-197.

[7] Wakeling, J.M., Uehli, K., Rozitis, A.I. (2006). Muscle fibre recruitment can respond to the mechanics of the muscle contraction. *Journal of the Royal Society Interface*, 3, 533-544.

[8] Wakeling, J., Delaney, R., Dudkiewicz, I. (2007). A method for quantifying dynamic muscle dysfunction in children and young adults with cerebral palsy. *Gait & Posture*, 25, 580-589.

[9] Weiderpass, H.A., Pachi, C.G.F., Yamamoto, J.F., Hamamoto, A., Onodera, A.N., Sacco, I.C.N. (2013). Time-frequency analysis methods for detecting effects of diabetic neuropathy. *International Journal for Numerical Methods in Biomedical Engineering*, 29, 1000-1010.

[10] Yang, J., Zhang, D., Frangi, A.F., Yang, J.Y. (2004). Two-dimensional PCA: A new approach to appearance-based face representation and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26, 131-137.

[11] Zhang, D.Q., Zhou, Z.H. (2005). (2D)(2)PCA: Two-directional two-dimensional PCA for efficient face representation and recognition. *Neurocomputing*, 69, 224-231.

[12] Kiatpanichagij, K., Afzulpurkar, N. (2009). Use of supervised discretization with PCA in wavelet packet transformation-based surface electromyrogram classification. *Biomedical Signal Processing and Control*, 4, 127-138.

[13] Jolliffe, I.T., Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of The Royal Society A – Mathematical Physical and Engineering Sciences*, 374, 20150202.

[14] Bro, R., Smilde, A.K. (2014). Principal component analysis. *Analytical Methods*, 6, 2812-2831.

[15] Englehart, K., Hudginsa, B., Chan, A.D.C. (2003). Continuous multifunction myoelectric control using pattern recognition. *Technology and Disability*, 15, 95-103.

[16] Huang, H., Xie, H.B., Guo, J.Y., Chen, H.J. (2012). Ant colony optimization-based feature selection method for surface electromyography signals classification. *Computers in Biology and Medicine*, 42, 30-38.

[17] Hakonen, M., Piitulainen, H., Visala, A. (2015). Current state of digital signal processing in myoelectric interfaces and related applications. *Biomedical Signal Processing and Control*, 18, 334-359.

[18] Kilby, J., Prasad, K., Mawston, G. (2016). Multi-channel surface electromyography electrodes: A review. *IEEE Sensors Journal*, 16, 5510-5519.

[19] Lucas, M.F., Gaufriau, A., Pascual, S., Doncarli, C., Farina, D. (2008). Multi-channel surface EMG classification using support vector machines and signal-based wavelet optimization. *Biomedical Signal Processing and Control*, 3, 169-174.

[20] Farina, D., do Nascimento, O.F., Lucas, M.F., Doncarli, C. (2007). Optimization of wavelets for classification of movement-related cortical potentials generated by variation of force-related parameters. *Journal of Neuroscience Methods*, 162, 357-363.

Received December 17, 2016.
Accepted May 15, 2017.

MEASUREMENT SCIENCE REVIEW, **17**, (2017), No. 3, 125-134

DE GRUYTER OPEN

MEASUREMENT SCIENCE REVIEW

Journal homepage: http://www.degruyter.com/view/j/msr

ISSN 1335-8871

The Application of Vibration Accelerations in the Assessment of Average Friction Coefficient of a Railway Brake Disc

Wojciech Sawczuk

Poznan University of Technology, Institute of Combustion Engines and Transport, Faculty of Machines and Transport, pl. M. Sklodowskiej-Curie 5, 60-965 Poznan, Poland, wojciech.sawczuk@put.poznan.pl

Due to their wide range of friction characteristics resulting from the application of different friction materials and good heat dissipation conditions, railway disc brakes have long replaced block brakes in many rail vehicles. A block brake still remains in use, however, in low speed cargo trains. The paper presents the assessment of the braking process through the analysis of vibrations generated by the components of the brake system during braking. It presents a possibility of a wider application of vibroacoustic diagnostics (VA), which aside from the assessment of technical conditions (wear of brake pads) also enables the determination of the changes of the average friction coefficient as a function of the braking onset speed. Vibration signals of XYZ were measured and analyzed. The analysis of the results has shown that there is a relation between the values of the point measures and the wear of the brake pads.

Keywords: Vibration, signal, diagnostic, railway disc brake, coefficient of friction.

1. INTRODUCTION

The growing speeds of trains force the application of friction disc brakes as the main train stopping systems. This is particularly the case for trains comprising a locomotive and passenger's coaches. In the case of traction sets, commonly referred to in Polish as 'EZT' a friction disc brake cooperates with an electrodynamic brake, in which the traction motors, acting as generators of additional resistance (regenerative braking) absorb the braking energy in the first stage of the process. Only in the second stage of braking are the disc brakes engaged at the speed of approx. 10 km/h following the drop in the efficiency of the electrodynamic brake. Due to the use of two braking systems, the work on vibroacoustic diagnostics (VA) of traction motors is presented in the works [10], [11], [12] or friction discs described in the works [28], [30].

The main advantages of a disc brake system include a stable course of the average friction coefficient as a function of speed of the onset of braking on the level of 0.35 or 0.37 as per [17] depending on the maximum train speed (0.37 for v_{max}=300 km/h, 0.35 for v_{max}=200 km/h) and good conditions of dissipation of the accumulated thermal energy to the environment, which is particularly efficient in the case of ventilated discs. A stable course of the braking process and the high value of average friction coefficient of as much as 0.4, when sintered brake pads are applied, as well as a quick transfer and dissipation of thermal energy enables the application of greater forces of the brake pads exerted on the disc, hence better braking power.

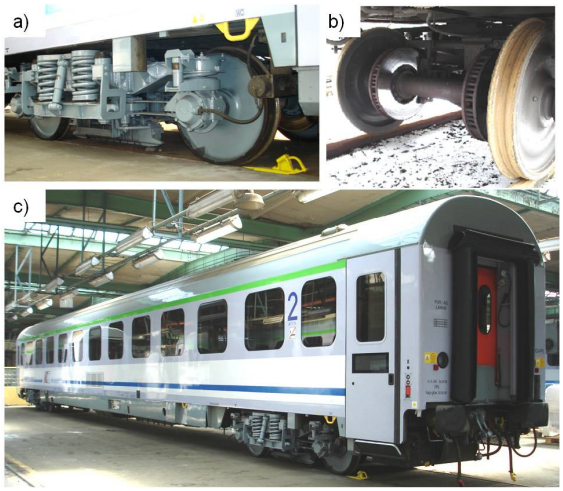

Fig.1. View of the passenger coach fitted with a disc brake: a) view of the MD 523 bogie, b) view of the wheelsets with the brake discs, c) view of the 136Amg passenger coach.

One of the few downsides of a disc brake system is that it is impossible to control the wear level of the components of the friction pair in operation. This is particularly conspicuous in passenger coaches, where the brake discs are fitted on the axles between the wheels as shown in Fig.1.b). In order to check the wear level of the friction pads, it is necessary to send the coach for service procedures so that the technicians can ascertain the wear from under the car and, if boundary wear is reached, renew the brake pads. In

DOI: 10.1515/msr-2017-0016

the automotive industry, the problem of brake pad wear has been resolved by applying a wear sensor in the brake pad. At the pad thickness of 2-3 mm the abrasion of the metal part of the sensor closes the circuit by connecting the sensor with the metal brake disc.

The selection of proper materials for the friction pair of a disc brake allows a wide range of the brake's friction characteristics, which directly influences the braking process. Instability of the braking process is an effect of a variety of phenomena discussed by many researchers. One of them is self-induced vibrations generated by the brake systems in different phases of the braking process, as presented in [7], [18], [27], [35]. The fluctuation of the instantaneous friction coefficient directly influences the vibration of variable amplitude. Another phenomenon of a thermal nature are the hot spots occurring on the brake disc, the effect of which is an uneven distribution of temperature and high local increments of temperature reaching 8, presented in [36] as a relation between the maximum temperature (at the hot spot) of the friction ring and the minimum temperature. The problem of the hot spots phenomenon on both the brake discs and clutch discs in motor vehicles has been presented in [2], [14], [16], [19], [24]. A separate problem adversely influencing the braking process are the chemical changes occurring in the brake pad material at high temperatures (above 200°C). This is related to the degassing of resins connecting the organic components of the metal brake pads, which results in the formation of a gas cushion between the brake pad and the disc. In literature, this effect is referred to as fading and has been described in [6], [8], [22]. In order to minimize the fading phenomenon, the manufacturers of the friction materials subject the brake pads to thermal processing at the temperature of 1200°C (scorching). Simultaneously, works are underway related to the materials for brake discs aiming at selecting appropriate components and thermal-chemical processing of irons, both gray and spheroid described in [1], [4], [5], [23], [25].

The aim of the investigations presented in the paper is an attempt to use the vibrations generated by the brake system while braking at different speeds to assess the braking process understood as change in the friction coefficient.

2. SELECTED MODELS OF VIBRATIONS IN DISC BRAKES

The initial models assumed that self-induced vibrations of the brake were related to the drop in the friction coefficient and an increase in the slide velocity. This is the case for a variety of materials but only for a limited range of speed change. If we assume a brake model of one degree of freedom as in Fig.2.b) then the equation of motion is expressed by the relation (1) [34].

$$m\ddot{x} + c\dot{x} + kx + N(\mu_{st} - \mu_o) \tag{1}$$

where: m_1, m_2 – masses of bodies,
N – pressing force,
k – coefficient of resilience,
c – coefficient of viscous damping,
v_o – velocity,
μ_{st} – coefficient of static friction,
μ_o – coefficient of kinetic friction.

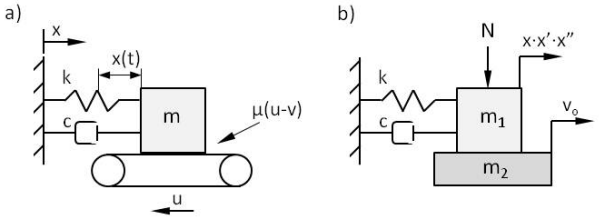

Fig.2. A model of vibrations of one degree of freedom: a) a model of a resilient friction system based on a conveyor; b) a model of mechanical system with feedback [34].

The system described in the equation of motion (1) will go into unstable vibrations depending on the values of the coefficient of damping μ, as shown in the relation (2).

$$c > \frac{N\Delta\mu}{2\Pi\sqrt{kmv_o^2}} \tag{2}$$

The first model that presents the possibility of instability of the friction system for a constant value of the friction coefficient was presented by Spurr as a slider with an angular support, as shown in Fig.3.a). This solution was improved by Jarvis and Earles. This mode, however, is geometrically inconsistent with the actual friction disc brake. Then, the condition of instability is determined with the relation (3) [34]:

$$\frac{1}{2}(\mu - tg\Theta\sin)2\Theta > \frac{C_p}{C_d} \tag{3}$$

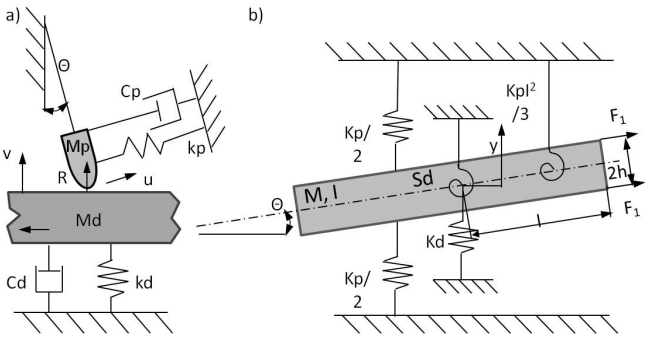

Fig.3. A model of instability of the friction system: a) disc - slider on an angular support; b) binary flutter of lumped parameters [34].

Another model was proposed by North and then Millner. This was a binary flutter model close to the model of a disc brake. The vibration mechanism in this case is similar to that of the fluttering effect of the wings of a plane. The race of the disc brake was replaced by a straight rigid beam of two degrees of freedom. The cross-section of the pad-race-pad system of the length of $2xl$ is shown in Fig.3.b). The condition of instability of the presented model is expressed by the relation (4) [34]:

$$\frac{8MIN \cdot \mu^2 h}{\left(I - \frac{1}{3}Ml^2\right)^2} > K_p > 0 \qquad (4)$$

where: $2h$ – thickness of the beam,
M – mass,
I – inertia,
$2l$ – length of the beam,
Kp – transverse rigidity of the pad,
$K_p l^2/3$ – rotary rigidity for the 2l portion.

As Crolla and Lang have proved, this and the other models do not entirely reflect the actual brake operation. Thanks to these models, however, we can obtain qualitative information needed in the design process of brake systems and in the search for design solutions eliminating some classes of vibrations and the resultant noise [34].

3. METHODOLOGY AND OBJECT OF RESEARCH

Research of a friction-vibroacoustic nature was performed on a certified inertia test stand at TABOR Institute of Rail Vehicles in Poznań. The test stand enables investigations of actual brake systems of rail vehicles (including pad and rail brakes) in the scale 1:1. Data obtained during tests on the certified brake inertia test stand are used for modeling the braking process or forecasting damage to brake system components. [9].

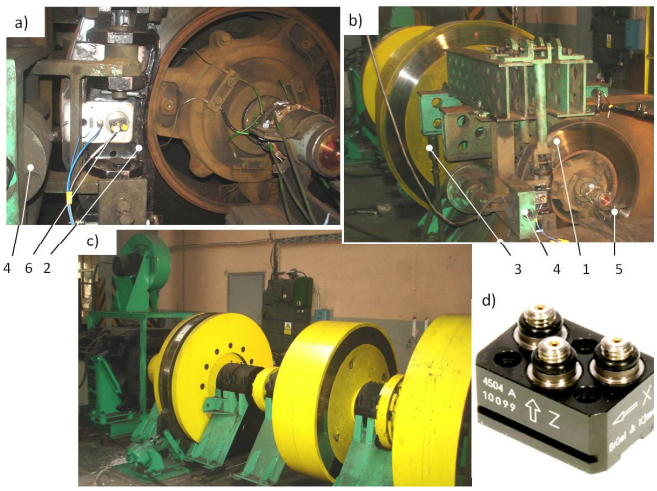

Fig.4. View of the research object on the disc brake test stand: a) view of the friction pair – brake disc and pads; b) View of the measurement section of the test stand; c) view of the drive component of the test stand with the rotating masses; d) View of the vibration transducer Brüel&Kjær type 4504A, 1 – brake disc, 2 – brake clasps and the pads, 3 – static force sensor referred to the brake radius, 4 – sensors of the pressing force of the pads to the brake disc, 5 – rotational speed sensor, 6 – vibration transducer.

The investigations were performed according to the assumptions of the active experiment described in [20], [21], the input parameters were intentionally modified in order to record the output parameters. The input parameters were the simulated speed of the onset of braking v, the pressing force of the pads on the brake disc N, the mass to be decelerated M, and the wear of the brake pads G. The output parameters were instantaneous tangential force F_t referred to the radius of braking r, instantaneous pressing force on the brake disc F_b, and instantaneous value of the vibration accelerations a. The observations of the input parameters on the changes of the output parameters were performed.

Fig.4. presents the brake test stand and the measurement equipment. For the measurements, the following sensors have been applied: HBM sensors (force) and B&K transducers (vibrations). The research object was a 590×110 railway brake disc with ventilating vanes and three sets of 175 FR20H.2 brake pads (a set of new pads of the thickness of 35 mm and two worn sets of 25 mm and 15 mm). During the tests a procedure of active experiment was applied, in which the output parameters were recorded following a modification of the input parameters.

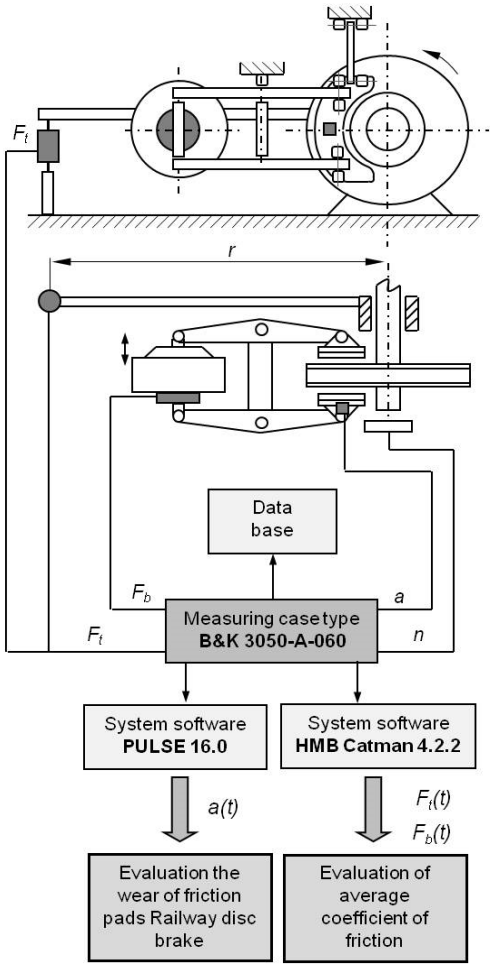

Fig.5. Diagram of the measurement track used in the investigations.

Tests of a tribological nature were performed according to the requirements stipulated in [17], [26]. During the tests, braking was performed from the speeds of 50, 80, 120, 160 and 200 km/h. During the tests, the pressing force of the pads to the disc of 25 kN was applied and the simulated mass to be decelerated was 5.7 tons.

The vibroacoustic tests were performed simultaneously with the friction (tribological) tests. One of the clasps was fitted with the vibration transducer shown in Fig.4.a). The measurements of the vibration accelerations were performed perpendicularly to the surface of the disc based on the experience of other researchers presented in [28], [32]. Fig.5. presents the diagram of the measurement track on the test stand additionally extended by the measurement of the vibration accelerations.

4. ANALYSIS OF THE RESULTS

In the domain of amplitudes of the analysis of vibration accelerations, the most frequently used are point measures that characterize a given vibration process with a single value according to [37]. Then, in vibroacoustic diagnostics in particular (VA), it is possible to determine the changes in the VA signal that result from a change of the technical condition of the tested object. Literature presents a variety of papers depicting the application of vibroacoustic diagnostics in road vehicles, rail vehicles, and aircrafts [3], [32], [33]. In order to determine the relation between the average friction coefficient and the vibrations generated by the brake system, in the first place it was proved that there exists a relation between the vibrations measured on the clasp and the technical condition of the system understood as the wear of the brake pads. To this end, for all analyzed speeds on the test stand, vibrations of the clasp were recorded, as shown in Fig.6. Then the following point measures were determined according to [37]:

1) RMS value of the vibration accelerations described with the relation (5):

$$A_{RMS} = \sqrt{\frac{1}{T}\int_0^T [a(t)]^2\, dt} \tag{5}$$

where: T – averaging time in [s],
$a(t)$ – instantaneous value of the vibration accelerations in [m/s^2].

2) The average values of the vibration accelerations described with the relation (6):

$$A_{AVERAGE} = \frac{1}{T}\int_0^t |a(t)|dt \tag{6}$$

3) Root value of the vibration acceleration described with the relation (7):

$$A_{SQUARE} = \left[\frac{1}{T}\int_0^T |a(t)|^{\frac{1}{2}}\, dt\right]^2 \tag{7}$$

4) Peak value of the vibration accelerations described with the relation (8):

$$A_{PEAK} = \left[\frac{1}{T}\int_0^T |s(t)|^{\infty}\, dt\right]^{\frac{1}{n}} \quad dla \quad n \to \infty \tag{8}$$

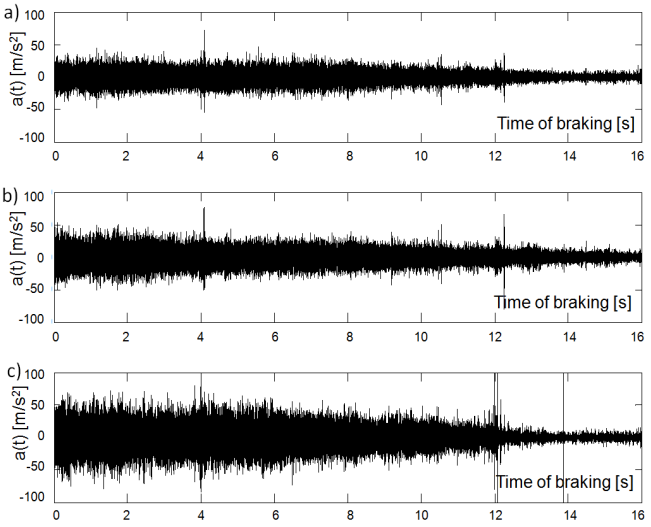

Fig.6. Instantaneous value of the vibration accelerations of the brake clasp with the brake pads when braking from the speed of 120 to the speed of 90 km/h with: a) new pads G1=35 mm; b) pads worn to 25 mm; c) pads worn to 15 mm.

For each point measure, the potential of using it as a parameter indicating the condition of the brake pads was verified. To this end, the coefficient of dynamics of the change of the diagnostic parameter was applied as per the relation (9) [13]:

$$D = 20\lg\left(\frac{A_2}{A_1}\right) \tag{9}$$

where: A_1 – value of the point measure (A_{RMS} or $A_{AVERAGE}$) determined when braking with new brake pads, G1=35 mm w [m/s^2],
A_2 – value of the point measure (A_{RMS} or $A_{AVERAGE}$) determined when braking with used brake pads, G2=25 and G3=15 mm w [m/s^2].

The recorded vibration signals were analyzed in Matlab 2014a. This included a deletion of fragments of signals not related to the braking process (startup of the test stand and coasting needed for the disc cooling), averaging as per relations (5)-(8) and determining of the dynamics of the change of the diagnostic parameter.

Table 1. presents the values of the point measures determined in the braking process for different speeds of the onset of braking.

The analysis of the results has shown that there is a relation between the values of the point measures and the wear of the brake pads (measuring the vibrations on the clasps with the brake pads fitted). Only in the case of low braking onset speeds (to 50 km/h) the dynamics of changes of the most heavily worn brake pad (15 mm) does not reach

6 dB. The braking processes for the speed starting from 80 km/h result in the peak value of the vibration acceleration and dynamics of the changes of the diagnostic parameter exceeding 6 dB. Higher braking onset speeds (160 and 200 km/h) cause also this point measure (A_{PEAK}) to find application in the measurements of the brake system vibrations.

Table 1. Values of the point measures for the braking process leading to a complete halt.

Point measure	Value of the point measure in m/s^2			Dynamics of changes in dB	
	Brake pad of the thickness of G1=35 mm	Brake pad of the thickness of G2=25 mm	Brake pad of the thickness of G3=15 mm	G2/ G1	G3/ G1
Speed of the onset of braking v=50 km/h					
A_{RMS}	6.88	9.36	10.63	2.67	3.78
$A_{AVERAGE}$	5.18	7.01	8.07	2.63	3.86
A_{SQUARE}	4.25	6.26	6.88	3.35	4.17
A_{PEAK}	45.06	59.25	65.35	2.38	3.23
Speed of the onset of braking v=80 km/h					
A_{RMS}	7.29	12.47	15.05	4.66	6.29
$A_{AVERAGE}$	5.55	9.39	11.48	4.57	6.32
A_{SQUARE}	11.10	20.44	25.06	5.30	7.08
A_{PEAK}	52.57	71.69	90.75	2.69	4.74
Speed of the onset of braking v=120 km/h					
A_{RMS}	8.65	13.95	17.68	4.15	6.21
$A_{AVERAGE}$	6.61	10.62	13.60	4.12	6.26
A_{SQUARE}	28.66	53.51	68.67	5.42	7.59
A_{PEAK}	59.63	95.22	111.55	4.07	5.44
Speed of the onset of braking v=160 km/h					
A_{RMS}	10.03	14.17	37.46	3.00	11.4
$A_{AVERAGE}$	7.59	10.67	18.66	2.96	7.8
A_{SQUARE}	62.34	106.42	168.47	4.64	8.6
A_{PEAK}	77.05	124.80	471.50	4.19	15.7
Speed of the onset of braking v=200 km/h					
A_{RMS}	9.68	12.57	96.31	2.27	19.7
$A_{AVERAGE}$	7.40	9.59	39.12	2.26	14.5
A_{SQUARE}	104.5	179.86	493.99	4.72	13.5
A_{PEAK}	85.20	131.30	918.63	3.76	20.6

Fig.7. presents a graphical relation between the analyzed point measures and the braking onset speeds, i.e. 50, 80, 120, 160 and 200 km/h. Fig.8. shows the relation between the dynamics of changes of the diagnostic parameter and the analyzed braking onset speeds. Following the information in Table 1., Fig.7. and Fig.8., it has been confirmed that the difference in the values of the point measures (between the new brake pads and the used ones) increases with the growing braking onset speed. Due to the dynamic nature of the braking process, particularly at high braking onset speeds, the vibrations generated by the brake system provide information related to the wear of the brake pads, which, in the further part of the paper will also be used in the assessment of the braking process understood as changes in the average friction coefficient.

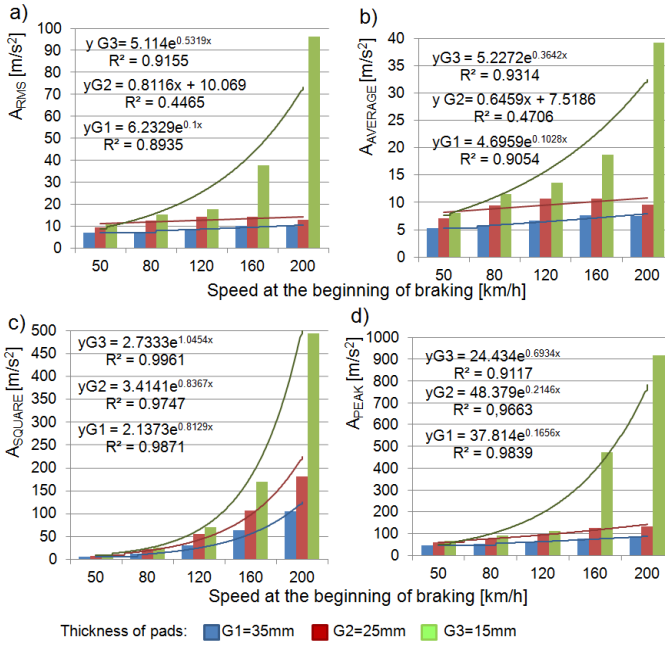

Fig.7. The relation of the values of selected point measures of vibration accelerations for three values of the brake pad thickness as a function of the braking onset speed: a) RMS value; b) average value; c) root value; d) peak value.

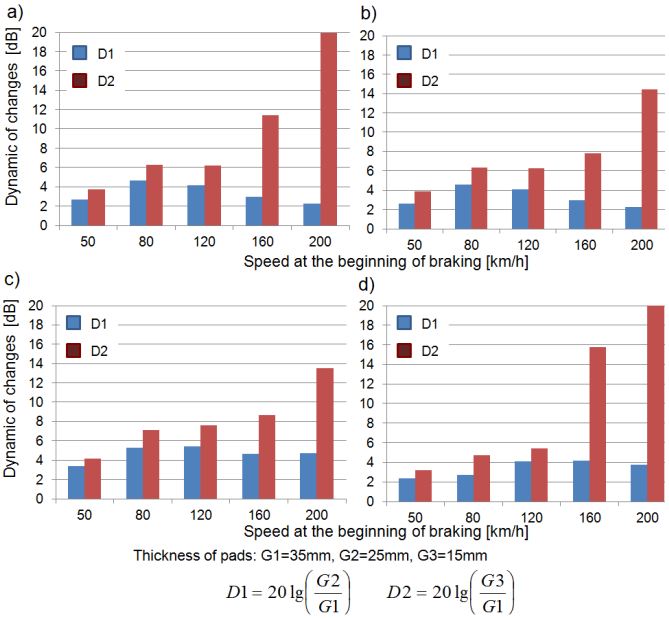

Fig.8. The relation of the dynamics of changes for selected point measures of vibration accelerations for three values of the brake pad thickness as a function of the braking onset speed: a) RMS value; b) average value; c) root value; d) peak value.

For further analyses, the effective and the average values of the vibration accelerations from the point measures were utilized due to the high value of the coefficient of dynamics of the changes of the diagnostic parameter for all braking onset speeds under analysis. In order to assess the brake pad wear, a reverse function to the approximating functions (shown in Fig.7.) was applied, so that it was possible to

assess the thickness of the brake pads based on the value of the point measures, averaging the entire process of braking until a full halt.

Fig.9. and Fig.10. show the relation between the brake pad thickness and a given point measure, i.e. the effective value and the average value.

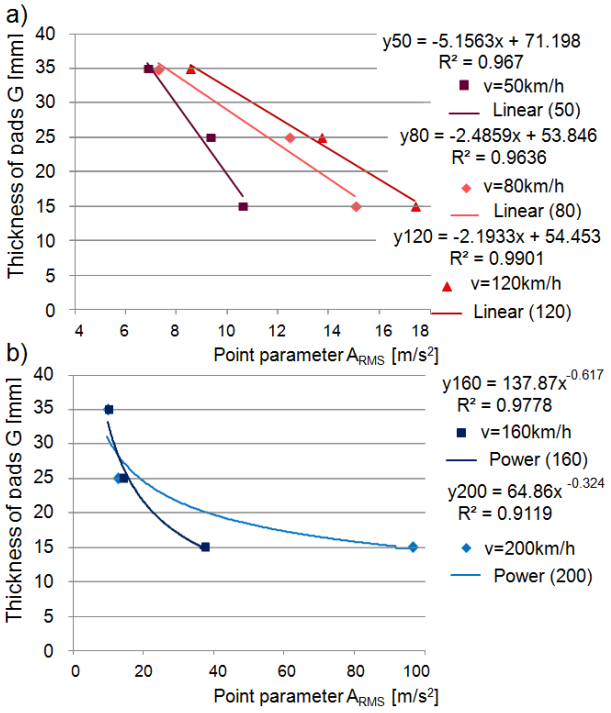

Fig.9. Dependence of the brake pad thickness on the RMS value of the vibration acceleration for the braking onset speeds of: a) v=50, 80 and 120 km/h; b) v=160 and 200 km/h.

The dependence of the brake pad thickness on the values of the point measures, due to the highest value of the coefficient of determinance R^2, was approximated with linear functions for the braking onset speeds from 50 to 120 km/h and exponential functions for the speeds of 160 and 200 km/h, which is expressed by the relations (10)-(19):

$$G_{(v=50)} = -5.1 \cdot A_{RMS\,(v=50)} + 71.2 \tag{10}$$

$$G_{(v=80)} = -2.5 \cdot A_{RMS\,(v=80)} + 53.8 \tag{11}$$

$$G_{(v=120)} = -2.2 \cdot A_{RMS\,(v=120)} + 54.4 \tag{12}$$

$$G_{(v=160)} = 137.9 \cdot A_{RMS(v=160)}^{-0.617} \tag{13}$$

$$G_{(v=200)} = 64.9 \cdot A_{RMS(v=200)}^{-0.324} \tag{14}$$

$$G_{(v=50)} = -6.7 \cdot A_{AVERAGE\,(v=50)} + 70.6 \tag{15}$$

$$G_{(v=80)} = -3.3 \cdot A_{AVERAGE\,(v=80)} + 53.8 \tag{16}$$

$$G_{(v=120)} = -2.8 \cdot A_{AVERAGE\,(v=120)} + 54.2 \tag{17}$$

$$G_{(v=160)} = 233.7 \cdot A_{AVERAGE\,(v=160)}^{-0.94} \tag{18}$$

$$G_{(v=200)} = 79.6 \cdot A_{AVERAGE(v=200)}^{-0.46} \tag{19}$$

In the further stage of the investigations carried out simultaneously with the vibration research, the authors measured instantaneous tangential force F_t referred to the braking radius and the instantaneous clamping force F_b of the friction pads to the brake disc in order to calculate the friction coefficient according to the relation (20) [17]:

$$\mu_a = \frac{F_t}{F_b} \tag{20}$$

Then, the average friction coefficient for all braking onset speeds under analysis was calculated as an integral of the instantaneous friction coefficient on the braking distance s_2 as per the relation (21) [17]:

$$\mu_m = \frac{1}{s_2} \int_0^{s_2} \mu \, ds \tag{21}$$

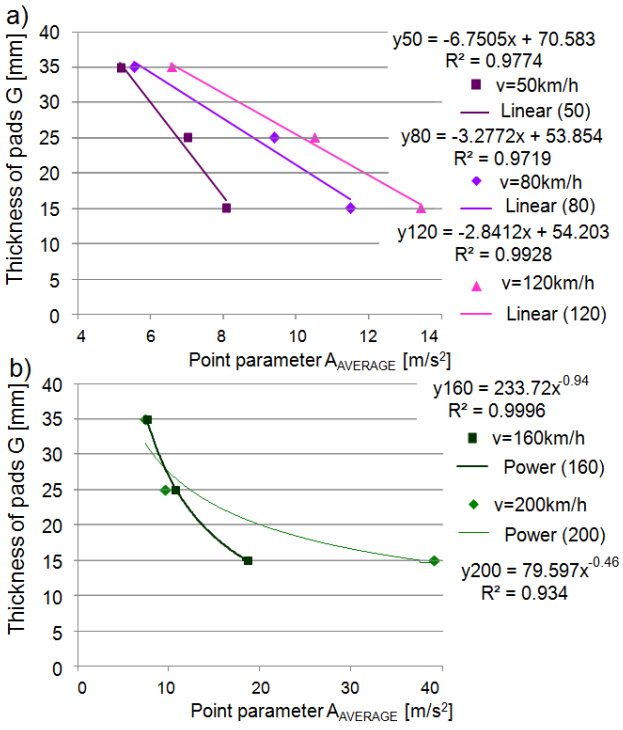

Fig.10. Dependence of the brake pad thickness on the average value of the vibration acceleration for the braking onset speeds of: a) v=50, 80 and 120 km/h; b) v=160 and 200 km/h.

Fig.11. shows the average friction coefficient as a function of the braking onset speed for three values of the brake pad thickness.

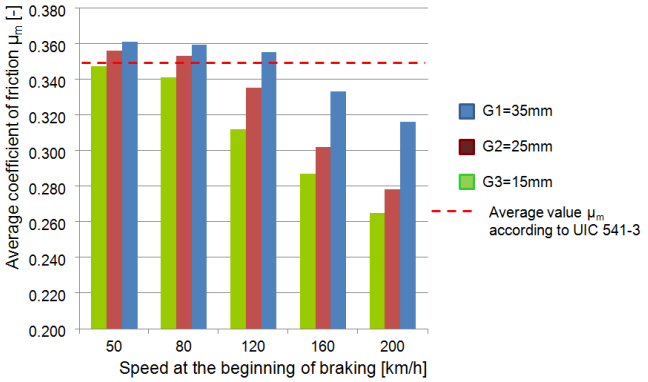

Fig.11. Dependence of the average friction coefficient μ_m on the braking onset speed at N=25 kN, M=5.7 t for three values of the brake pad thickness.

Analyzing Fig.11., it can also be observed that the average friction coefficient also depends on the conditions of the brake pads (their wear). As the brake pads wear down, the value of the average friction coefficient decreases. Fig.12. and Fig.13. show the dependence of μ_m on the thickness of the brake pads. Additionally, the average friction coefficient was approximated with the linear function as shown by the relations presented in the graphs.

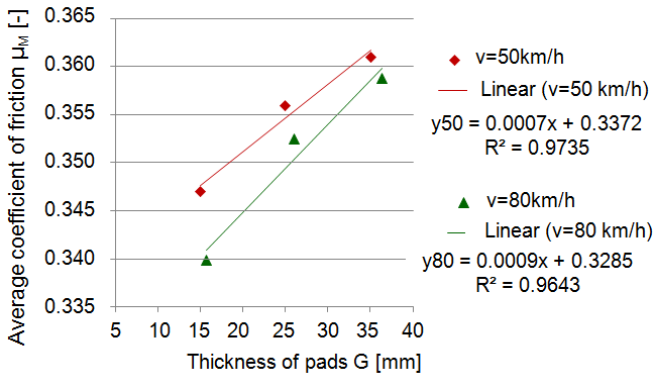

Fig.12. Dependence of the average friction coefficient μ_m on the thickness of the brake pads for the braking onset speeds of v=50 and 80 km/h.

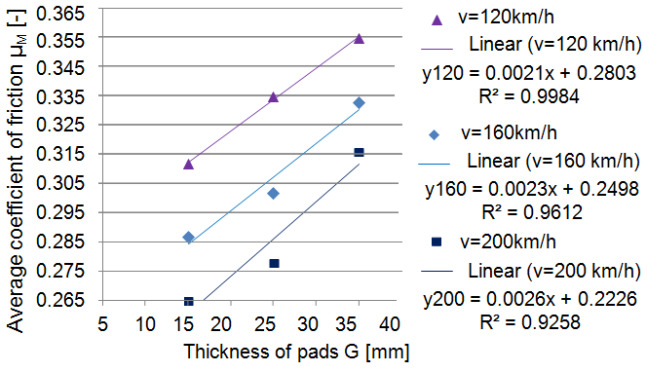

Fig.13. Dependence of the average friction coefficient μ_m on the thickness of the brake pads for the braking onset speeds of v=120, 160 and 200 km/h.

Because of the occurrence of linear (low braking onset speeds) or non-linear (higher braking onset speeds) relation between the values of the brake pad thickness and the values of the point measures and a linear relation between the average friction coefficient and the brake pad thickness, a relation aiming at the estimation of the average friction coefficient based on the recorded vibrations was determined through the method of substitution of two functions. A general form of the determination of the average value of the friction coefficient is expressed by equations (22) and (23). Equation (22) is based on two linear functions and equation (23) is a result of joining a non-linear function with a linear one.

$$\begin{cases} G = a_1 A_{RMS}(or\ A_{AVERAGE}) + b_1 \\ \mu_m = a_a G + b_2 \\ \mu_m = a_1 a_2 A_{RMS}(or\ A_{AVERAGE}) + a_2 b_1 + b_2 \\ for\ v = 50, 80, 120 \left[\frac{km}{h}\right] \end{cases} \tag{22}$$

$$\begin{cases} G = a_1 A_{RMS}(or\ A_{AVERAGE})^{b_1} \\ \mu_m = a_a G + b_2 \\ \mu_m = a_1 a_2 A_{RMS}(or\ A_{AVERAGE})^{b_1} + b_2 \\ for\ v = 160, 200 \left[\frac{km}{h}\right] \end{cases} \tag{23}$$

Applying the relations pertaining to the brake pad wear as a function of the values of the point measures of the vibration accelerations and between the average friction coefficient and the brake pad wear based on the assumptions of functions (22) and (23), the following relations for the assessment of the average friction coefficient were determined:

$$\mu_{m,(v=50)} = -0.0036 \cdot A_{RMS(v=50)} + 0.387 \tag{24}$$

$$\mu_{m,(v=80)} = -0.0022 \cdot A_{RMS(v=80)} + 0.376 \tag{25}$$

$$\mu_{m,(v=120)} = -0.0046 \cdot A_{RMS(v=120)} + 0.394 \tag{26}$$

$$\mu_{m,(v=160)} = 0.331 \cdot A_{RMS(v=160)}^{-0.617} + 0.249 \tag{27}$$

$$\mu_{m,(v=200)} = 0.168 \cdot A_{RMS(v=200)}^{-0.324} + 0.222 \tag{28}$$

$$\mu_{m,(v=50)} = -0.0047 \cdot A_{AVERAGE(v=50)} + 0.386 \tag{29}$$

$$\mu_{m,(v=80)} = -0.0029 \cdot A_{AVERAGE(v=80)} + 0.376 \tag{30}$$

$$\mu_{m,(v=120)} = -0.0059 \cdot A_{AVERAGE(v=120)} + 0.394 \tag{31}$$

$$\mu_{m,(v=160)} = 0.56 \cdot A_{AVERAGE(v=160)}^{-0.94} + 0.249 \tag{32}$$

$$\mu_{m_{\cdot(v=200)}} = 0.21 \cdot A_{AVERAGE(v=200)}^{-0.46} + 0.222 \quad (33)$$

where: μ_m – average value of the friction coefficient [-],
A_2 – value of the effective point A_{RMS} or average $A_{AVERAGE}$ measure in [m/s^2].

The analyses of the results in the domain of amplitudes have shown that, based on the point measures of vibration accelerations of the brake system discussed in the paper, a diagnostic of the brake conditions is possible as described earlier in [15], [28], [31], including the assessment of the braking process through the determination of the average friction coefficient. Table 2. and Table 3. present the relative percentage error of the suitability of the model of the average friction coefficient based on functions (23)-(32) against the values determined during the tests on a certified railway brake test stand. The authors relied on the two point measures of the vibration accelerations measured on the brake clasps during the braking process. It should be noted that also in the field of vibroacoustic diagnostics of the braking systems are used more advanced analyses as frequency analysis presented in the work [29] or time-frequency analysis [30]. However, in spite of the greater accuracy of the friction wear estimation, in the case of a diagnostic system, such analyses will generate a more sophisticated vibroacoustic signal processing system.

Table 2. Relative percentage error of the suitability of the model of the average friction coefficient to the test results based on the determined effective values A_{RMS} of the vibration accelerations.

For the point measure A_{RMS}			
Speed [km/h]	For a new brake pad G1	For the worn brake pad G2	For the worn brake pad G3
v=50	0.3	0.8	0.5
v=80	0.3	1.3	0.6
v=120	0.2	1.5	0.2
v=160	1.3	3.7	0.9
v=200	4.3	6.1	1.8

Table 3. Relative percentage error of the suitability of the model of the average friction coefficient to the test results based on the determined average values $A_{AVERAGE}$ of the vibration accelerations.

For the point measure $A_{AVERAGE}$			
Speed [km/h]	For a new brake pad G1	For the worn brake pad G2	For the worn brake pad G3
v=50	0.2	0.8	0.3
v=80	0.3	1.2	0.5
v=120	0	1.1	0.6
v=160	0.2	2.4	0.8
v=200	3.3	6.2	1.6

The graphical representation of the suitability of the regressive model of the value estimation of the average friction coefficient against the test results on the railway brake test stand is shown in Fig.14.

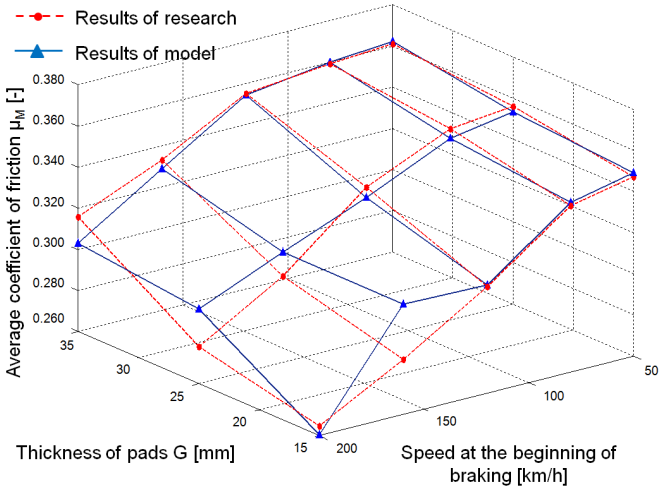

Fig.14. Relation between the average friction coefficient μ_m obtained during the investigations and the regressive model obtained from equations (23)-(32) as a function of the thickness of the brake pads and the braking onset speeds.

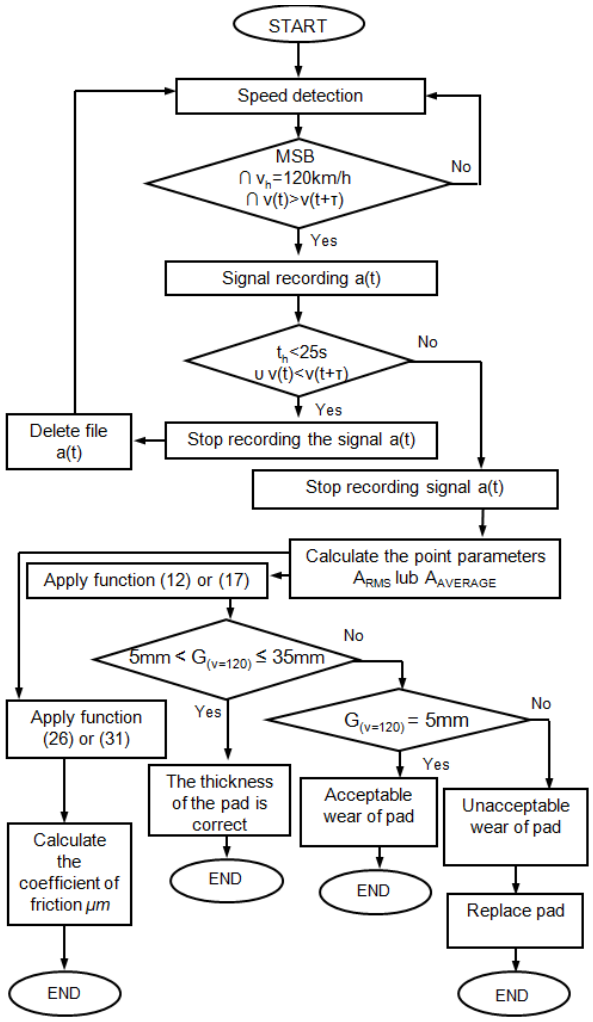

Fig.15. Algorithm for evaluation of wear of friction pads and assessment of braking process on example of braking from speed v=120 km/h, MSB - Make stopping braking, τ – time increment.

Fig.15. shows the algorithm for the selected braking speed for simultaneous evaluation of friction pad wear and the evaluation of braking. The braking process is evaluated by determining the mean friction coefficient. The algorithm used to process the vibroacoustic signal generated by the friction pad during braking is used to assess friction wear. The wear assessment is made in three ranges, i.e. for the thickness of the pad between 5 and 35mm, for a thickness of 5mm, and for a thickness of less than 5mm where a pad replacement message is generated for the new pad. Then, by applying the functions described in (24) - (33), it is possible to determine the mean coefficient of friction for the braking speed from 50 to 200 km/h.

7. CONCLUSION

The paper presents a possibility of a wider application of vibroacoustic diagnostics (VA), which, aside from the assessment of technical conditions (in this case, wear of the brake pads) also enables the determination of the changes of the average friction coefficient as a function of the braking onset speed. It results from the fact of a heavy dependence of the diagnostic parameter on the brake pad wear expressed with the dynamics of changes exceeding 6 dB as well as the dependence of the average friction coefficient on the speed and brake pad wear. The connection of both functions, i.e. the wear of the brake pads/values of the point measures of vibration accelerations and the average friction coefficient/wear of the brake pads allows determining the linear (low speeds) and non-linear (greater speeds) regressive models for the assessment of the average friction coefficient. In the analysis of vibration signals, point measures in the domain of amplitudes are sufficient being averaged to a single value of instantaneous changes of the vibration accelerations of a selected component of a brake system. An error in the reproduction of the model of average friction coefficient, based on the determined point measures during the braking process, reaches 6 % only for some braking onset speeds. The article presents an algorithm for simultaneous evaluation of wear of friction pads and evaluation of braking process based on linear functions (24)-(33). The application of more sophisticated analysis tools such as FFT and STFT frequency analysis with filters or the application of a time-frequency analysis will enable better compliance of the model results with the actual results, however, it will result in a more complicated signal processing system and ultimately a more complex diagnostic system.

ACKNOWLEDGMENTS

The project has been financed by the National Centre for Research and Development, program LIDER V, contract No. LIDER/022/359/L-5/13/NCBR/2014

REFERENCES

[1] Aranganathan, N., Bijwe, J. (2016). Development of copper-free eco-friendly brake-friction material using novel ingredients. *Wear*, 352-353, 79-91.

[2] Belhocine, A., Bouchetara, M. (2012). Thermomechanical modelling of dry contacts in automotive disc brake. *International Journal of Thermal Sciences*, 60, 161-170.

[3] Chen, J., Randall, R.B., Peeters, B. (2016). Advanced diagnostic system for piston slap faults in IC engines, based on the non-stationary characteristics of the vibration signal. *Mechanical System and Signal Processing*, 75, 434-454.

[4] Collignon, M., Regheere, G., Cristol, A.L., Desplanques, Y., Balloy, D. (2013). Braking performance and influence of microstructure of advanced cast irons for heavy goods vehicle brake discs. *Journal of Engineering Tribology*, 227 (8), 930-940.

[5] Crăciun, A. (2015). Evolution of materials for motor vehicles brake discs. *ANNALS of Faculty Engineering Hunedoara - International Jurnal of Engineering*, 13 (3), 149-154.

[6] Dubarry, M., Svoboda, V., Hwu, R., Liaw, B.Y. (2007). Capacity and power fading mechanism identification from a commercial cell evaluation. *Journal of Power Sources*, 165, 566-572.

[7] Fazio, O., Nacivet, S., Sinou, J. (2015). Reduction strategy for a brake system with local frictional non-linearities – Application for the prediction of unstable vibration modes. *Applied Acoustics*, 91, 12-24.

[8] Fidlin, A., Bäuerle, S., Boy, F. (2015). Modelling of the gas induced fading of organic linings in dry clutches. *Tribology International*, 92, 559-566.

[9] Gill, A., Kadziński, A. (2015). The determination procedure of the onset of the object wear-out period based on monitoring of the empirical failure intensity function. *Eksploatacja i Niezawodność – Maintenance and Reliability*, 17 (2), 282–287.

[10] Glowacz, A. (2016). Fault diagnostics of acoustic signals of loaded synchronous motor using SMOFS-25-EXPANDED and selected classifiers. *Tehnicki Vjesnik-Technical Gazette*, 23 (5), 1365-1372.

[11] Glowacz, A., Glowacz, Z. (2017). Diagnosis of stator faults of the single-phase induction motor using acoustic signals. *Applied Acoustics*, 117 (A), 20-27.

[12] Glowacz, A. (2016). Fault diagnostics of DC motor using acoustic signals and MSAF-RATIO30-EXPANDED. *Archives of Electrical Engineering*, 65 (4), 733-744.

[13] Gryboś, R. (2009). *Machine Vibrations*. Gliwice, Poland: Publishing House of Silesia University of Technology, 214.

[14] Grzes, P., Oliferuk, W., Adamowicz, A., Kochanowski, K., Wasilewski, P., Yevtushenko, A.A. (2016). The numerical-experimental scheme for the analysis of temperature field in a pad-disc braking system of a railway vehicle at single braking. *International Communications in Heat and Mass Transfer*, 75, 1-6.

[15] Hoguet, T. (2007). Brake tests at Bosch. *Brüel & Kjær Magazine*, 2, 22-24.

[16] Kasem, H., Brunel, J.F., Dufrénoy, P., Siroux, M., Desmet, B. (2011). Thermal levels and subsurface damage induced by the occurrence of hot spots during high-energy braking. *Wear*, 270, 355-364.

[17] *Kodeks UIC 541–3. Hdisc brake and its application. Conditions of brake pad permission for use*. 7th edition, June 2010, 10-24.

[18] Kruse, S., Tiedemann, M., Zeumer, B., Reuss, P., Hetzler, H., Hoffmann, N. (2015). The influence of joints of friction induced vibration in brake squeal. *Journal of Sound and Vibration*, 340, 239-252.

[19] Kumar, M., Boidin, X., Desplanques, Y., Bijwe, J. (2011). Influence of various metallic fillers in friction materials on hot-spot appearance during stop braking. *Wear*, 270, 371-381.

[20] Leszek, W. (2006). *Selected methodological issues of empirical research*. Institute of Maintenance Technology, Radom, Poland, 142-153.

[21] Mańczak, K. (1976). *Techniques of Experiment Planning*. Warsaw, Poland: WNT, 76-84.

[22] Neis, P.D., Ferreira, N.F., Lorini, F.J. (2011). Contribution to perform high temperature tests (fading) on a laboratory-scale tribometer. *Wear*, 271, 2660-2664.

[23] Paczkowska, M. (2016). The evaluation of the influence of laser treatment parameters on the type of thermal effects in the surface layer microstructure of gray irons. *Optics and Laser Technology*, 76, 143-148.

[24] Panier, S., Dufrénoy, P., Weichert, D. (2004). An experimental investigation of hot spots in railway disc brakes. *Wear*, 256, 764-773.

[25] Peveca, M., Oder, G., Potrč, I., Šraml, M. (2014). Elevated temperature low cycle fatigue of grey cast iron used for automotive brake discs. *Engineering Failure Analysis*, 42, 221-230.

[26] Polish Committee for Standardization. (2016). *Railway – brake discs of rail vehicles – Part 3: Brake discs, properties of a disc brake and the friction pair, classification.* PN-EN 14535-3, 12-16.

[27] Rudolph, M., Popp, K. (2001). Brake squeal. In *Detection, Utilization and Avoidance of Nonlinear Dynamical Effects in Engineering Applications: Final Report of a Joint Research Project Sponsored by the German Federal Ministry of Education and Research*. Shaker Verlag, 197–225.

[28] Sawczuk, W. (2016). Application of vibroacoustic diagnostics in the evaluation of wear of friction pads in a railway disc brake. *Maintenance and Reliability*, 18 (4), 565–571.

[29] Sawczuk, W. (2015). Application of selected frequency characteristics of vibration signal for the evaluation of the braking process for railway disc brake. *Diagnostyka - Applied Structural Health, Usage and Condition Monitoring*, 16 (3), 33-38.

[30] Sawczuk, W., Szymanski, G.M. (2016). Diagnostics of the railway friction disc brake based on the analysis of the vibration signals in terms of resonant frequency. *Archive of Applied Mechanics*, 86, 1-15.

[31] Segal, L. (1999). Diagnostic method for vehicle brake. *NDT&E International*, 32, 369-373.

[32] Szymanski, G.M., Josko, M., Tomaszewski, F., Filipiak, R. (2015). Application of time-frequency analysis to the evaluation of the condition of car suspension. *Mechanical System and Signal Processing*, 58-59, 298-308.

[33] Szymanski, G.M., Josko, M., Tomaszewski, F. (2016). Diagnostics of automatic compensators of valve clearance in combustion engine with the use of vibration signal. *Mechanical System and Signal Processing*, 68-69, 479-490.

[34] Ścieszka, S.F. (1998). *Friction Brakes. Material, Design and Tribological Issues*. Publishing House Gliwice-Radom, Poland, 106-110.

[35] Triches, M., Samir, N.Y., Jordan, R. (2008). Analysis of brake squeal noise using finite element method. A parametric study. *Applied Acoustics*, 69, 147-162.

[36] Wirth, X. (1998). Improving the performance of disc brakes on high-speed rail vehicles with a novel types of brake pad: Isobar. *RTR,* 1, 24-29.

[37] Żółtowski, B., Lukasiewicz, M. (2012). *Vibration Diagnostics of Machines*. WNT Radom, Poland, 113-115.

Received December 02, 2016.
Accepted May 16, 2017.

MEASUREMENT SCIENCE REVIEW, **17**, (2017), No. 3, 135-144

DE GRUYTER OPEN

MEASUREMENT SCIENCE REVIEW

Journal homepage: http://www.degruyter.com/view/j/msr

ISSN 1335-8871

Identification and Adjustment of Guide Rail Geometric Errors Based on BP Neural Network

Gaiyun He, Can Huang, Longzhen Guo, Guangming Sun, Dawei Zhang

Key Laboratory of Mechanism Theory and Equipment Design of Ministry of Education, Tianjin University, Tianjin 300072, China, hegaiyun@tju.edu.cn

The relative positions between the four slide blocks vary with the movement of the table due to the geometric errors of the guide rail. Consequently, the additional load on the slide blocks is increased. A new method of error measurement and identification by using a self-designed stress test plate was presented. BP neural network model was used to establish the mapping between the stress of key measurement points on the test plate and the displacements of slide blocks. By measuring the stress, the relative displacements of slide blocks were obtained, from which the geometric errors of the guide rails were converted. Firstly, the finite element model was built to find the key measurement points of the test plate. Then the BP neural network was trained by using the samples extracted from the finite element model. The stress at the key measurement points were taken as the input and the relative displacements of the slide blocks were taken as the output. Finally, the geometric errors of the two guide rails were obtained according to the measured stress. The results show that the maximum difference between the measured geometric errors and the output of BP neural network was 5 μm. Therefore, the correctness and feasibility of the method were verified.

Keywords: Guide rail geometric error, stress, the test plate, finite element model, BP neural network.

1. INTRODUCTION

The precision maintenance is one of the most important criteria of machine tools, which is affected by the stiffness, geometric errors and thermal errors [1]-[2]. After the machine tools assembly had been completed, the stress is gradually released with time. Thus, the precision of the machine tools cannot maintain a high level for a long time. The distribution of stress in machine tools is influenced by many factors, and the geometric error is one of the key factors.

The geometric errors of the guide rails mainly include the straightness and the parallelism between the two guide rails. In the manufacturing and installation process, linear rolling guide rail will inevitably produce errors, resulting in the relative position of four slider blocks changing with the table movement [3]. The stress generated within the guide system is thereby affecting the precision maintenance of the table.

Over the past decades, the geometric errors of guide rails and their influence on motion accuracy of the table have been studied. Mahdi Rahmani investigated the geometric accuracy and its effect on the performance of the guiding system in machine tools with finite element method [4]. Eiji Shamoto et al. used the transfer function model to establish the relationship between geometric errors of hydrostatic guide and table motion errors [5]. The geometric errors of the guide rail were calculated by measuring motion errors of the table. Gyungho Khim et al. took linear rolling guide as the subject and established the force balance equation of the table using the same method. The calculations were carried out via the Hertz contact theory. In addition, the influence of the guide rail geometric errors on the 5-DOF motion errors was analyzed [6]-[7]. Gyu Ha Kim proposed a new transfer function method based on reaction force-moment model with a double spring system. The influence of the pitch error of the slide block on the moment was also considered [8]. However, the deformation of the table and the slide block was not considered in their research. It was deemed that omitting the elasticity of the table and slide block was an excessive simplification.

At present, the relationship between the deformation field and the distribution of stress is rarely researched. Li et al. presented a test method of assembly stress based on the strain test [9]. The stress distribution of the machine tools was studied in a different assembly process. The results showed that the tightening sequence of bolts has little influence on the assembly stress. Paolo Bosetti presented a novel measurement system for measuring deformation field of machine tools components [10]. The struts in the system were instrumented with a strain sensor which provided their longitudinal strain values. An algorithm was used to evaluate the discrete displacements filed by calculating the

DOI: 10.1515/msr-2017-0017

node positions on the basis of the strut longitudinal deformations. Liu et al. employed optical fiber Bragg grating sensors to measure the strain field of machine tool components in real time [11]. The deformation field was obtained according to the structural dimension of the component. Consequently, the displacement of the tool tip was obtained.

The finite element method has been widely used in the engineering field. Edward Chlebus described a method of calculating the static performance of guide rails using the finite element analysis [12]. The effectiveness of the method of modeling and calculation was confirmed by experiments. Pawel Majda used the finite element method to study the relationship between kinematic straightness errors and angular errors of the table [13]. The simulation results showed that no close relationship was found between these two types of errors. Meanwhile, a method of analytical and experimental examinations to research the influence of the guide rail geometric errors on joint kinematic errors was proposed [14]. The results verified that the deformation of table could be a significant source of errors in volumetric error models. However, the relationship between the geometric errors of the guide rails and the stress distribution of the table has not been researched.

In this paper, a new measuring instrument and method were adopted to identify geometric errors of guide rails. The self-designed test plate was regarded as the elastic element. The stress state of the test plate was analyzed when the relative position of the slide blocks changed. The mapping relationship between relative displacements of slide blocks and the stress of the test plate were established by BP neural network. The measured stress of the test plate was input into the BP neural network. Consequently, the geometric errors of the two guide rails in X and Y directions can be obtained, which can provide the basis for the adjustment of the guide rail geometric errors.

2. THE STRUCTURE OF THE TEST PLATE

Due to the geometric errors of the guide rails, the relative displacements of sliders are produced during the movement. When the stiffness of the table is sufficient enough, the deformation occurs mainly on the slide blocks [15]. The additional load on the slide blocks was increased, which affects the useful life of the slide blocks and the precision maintenance of the table. When the stiffness of the table is insufficient, the table will deform greatly. And the stress state of the table will change with the relative displacements of the slide blocks. Because of the complex structure of the actual table, it is difficult to measure the stress. Therefore, a simplified test plate was designed as the stress measuring instrument. And the relative displacements of the slide blocks were reflected by the stress state of the test plate.

A. Structure of the test plate

According to the distance between the two guide rails, the length of the test plate l_m was designed to be 560 mm, the width l_w was 560 mm and the thickness l_t was 45 mm.

In order to highlight the stress in different directions on the surface of the test plate, the ribs were arranged in the form of the structure shown in Fig.1. Eight ribs were equally distributed on the test plate. The width of the ribs l_{rw} was 15 mm, the height l_{rh} was 25 mm. The angle α between the ribs was 45 °.

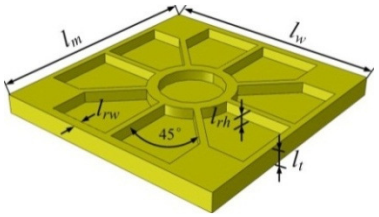

Fig.1. Basic dimensions of the test plate.

B. Stress comparison with different structures

In order to demonstrate the advantage of the test plate, we made a comparison with a flat plate. The flat plate has the same basic dimensions except that the thickness was 20 mm without ribs. The vertical displacement was applied to the contact surface between the test plate and the second slide block. The same vertical displacement was applied to the same position of the flat plate. In Fig.2., the stress nephogram of the two kinds of plates is compared. P_1, P_2, P_3, P_4 are the serial numbers of the slider blocks.

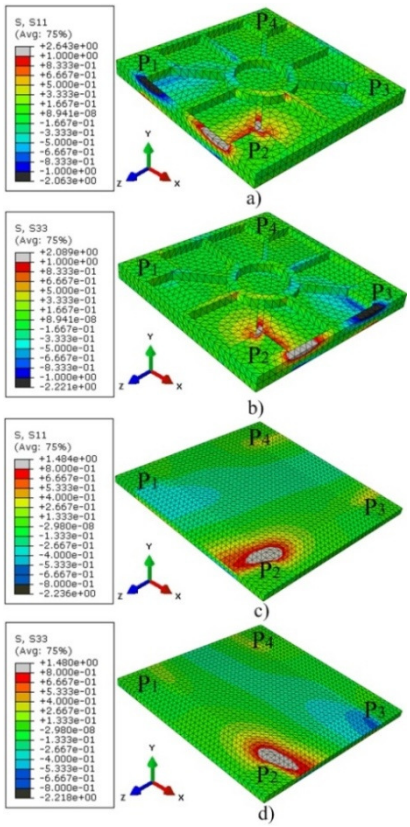

Fig.2. The stress nephogram of the plates when the vertical displacement was applied to the second contact surface. a) and b): the stress distribution of the test plate; c) and d): the stress distribution of the flat plate.

In Fig.2.a) and Fig.2.b), the stress is mainly distributed on the ribs and the direction of stress is consistent with the direction of the ribs. In Fig.2.c) and Fig.2.d), the stress in different directions is mainly distributed in a certain area. So, it is more convenient to measure the stress of the test plate. The stress state of the test plate will be changed with the relative displacements of the slide blocks, and then get the geometric errors of guide rails.

3. FINITE ELEMENT MODEL

In modeling, the finite element method (FEM) was used. The sub-assemblies were discretized with solid elements, which made it possible to allow the elasticity within the linear elastic range for guide rail, slide block, and test plate.

A. The roller contact stiffness

When the roller was subjected to a load, the elastic deformation occurred between the roller and the raceway. Considering the usefulness of the performed analytical examinations, it seemed that omitting the elasticity of single roller was an excessive simplification, which decreased the reliability of the performed analyses.

In the finite element model, the roller can be equivalent to the spring element [14]. Considering geometric non-linearity (only compression), the contact deformation of the single roller within two grooves can be calculated by the Palmgren empirical formula [16].

$$\delta = 1.36\frac{(\eta Q_n)^{0.9}}{(l_e)^{0.8}} \tag{1}$$

$$\eta = \frac{1}{E'} = \frac{1-v_1^2}{E_1} + \frac{1-v_2^2}{E_2} \tag{2}$$

where E_1, E_2 represent Young´s modulus of the roller and the raceway, respectively, MPa; v_1, v_2, the Poisson ratio of the roller and groove material; δ, deformation of a single roller; Q_n, force acting on a single roller, N; l_e, length of roller, mm; η, parameter that depends on E and v.

This research adopted the linear rolling guide produced by the company of THK. Parameters of the guide rail are shown in Table 1.

Table 1. Parameters of linear rolling guide [17].

Parameters	**Value**
Young´s modulus E_1, E_2 [GPa]	206
Poisson ratio v_1, v_2	0.3
Length of roller l_e [mm]	8
The diameter of the roller d [mm]	4
The number of single row roller Z	21

Substituting the parameters into (1), the relationship between the deformation δ of the roller and the load Q_n was obtained. The stiffness curve shown in Fig.3. is plotted. The roller contact stiffness K=2.77×10^5 N/mm was obtained by the linear fitting method. On the other hand, K is also the stiffness of the spring element in the finite element model.

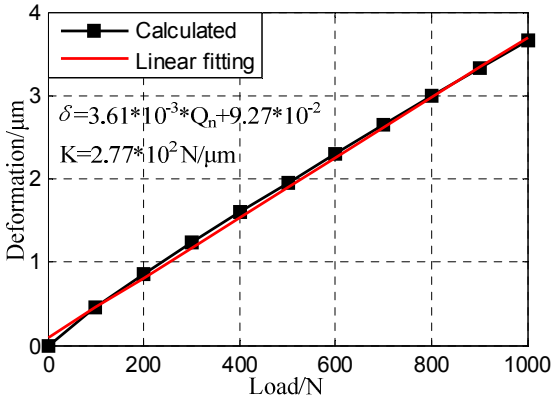

Fig.3. The stiffness curve of roller.

B. Finite element model of guide rail and slide block

A contact element was used to model the contact phenomena and the elasticity of a single roller. With regard to the physical modeling, each roller was equivalent to a spring element, which connects the slider with the guide rail in the finite element model. The idea of modeling is shown in Fig.4.

The roller contact stiffness calculated by (1) and (2) was assigned to the spring element. When Q_n was lower than zero, the gap occurred between the roller and raceway. Therefore, the elastic of spring element was zero in the finite element model. When Q_n was greater than zero, the deformation occurred between the roller and raceway. The elastic of spring element was K (K=2.77×10^5 N/mm) in the finite element model. The reference length of the spring was equal to the diameter of the roller. Generally, the diameter of the roller should be slightly larger than the normal spacing, so that the roller and the raceway surface will produce interference fit.

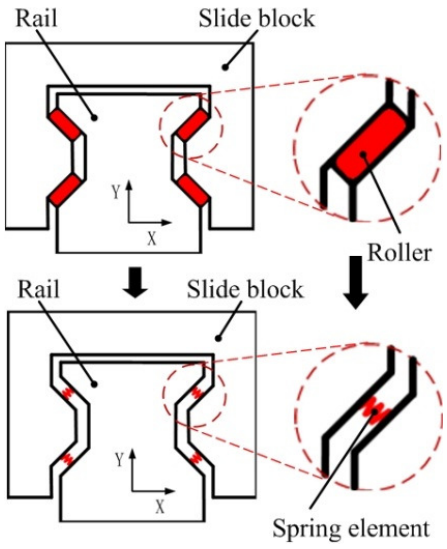

Fig.4. The roller was replaced by the spring element.

Necessary geometric dimensions and preload for the slider were taken from the product list [17]. Preload was 0.08*C*; *C* is the dynamic capacity and equals to 22.8 kN. Dividing the preload by the number of rollers, the load subjected to each roller could be obtained. Based on the characteristics described by (1), the deformation δ of each roller was 2 μm. It is possible to increase the spring element reference length by 2 μm to simulate the preload.

As shown in Fig.5., the coupling constraints were established between the end points of the spring element and the contact areas of the raceway. The coupling nodes shared the load. The Young´s modulus of guide rail and slider block were 206 GPa, the Poisson ratio was 0.3. The guide rail and slider block were meshed with hexahedron, and obtained a total of 120 122 elements and 133 738 nodes.

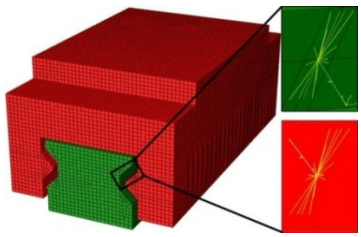

Fig.5. Finite element model of guide rail and slide block.

In the finite element model, the base of the guide rail was fixed. A load of 0 to 30 kN was applied to the upper surface and the side surface of the slide block, respectively. In order to avoid the stress concentration, the concentrated force was transformed into the pressure distributed on the surface. The displacements of slider in vertical and horizontal direction were extracted, so the stiffness curves of slider were obtained.

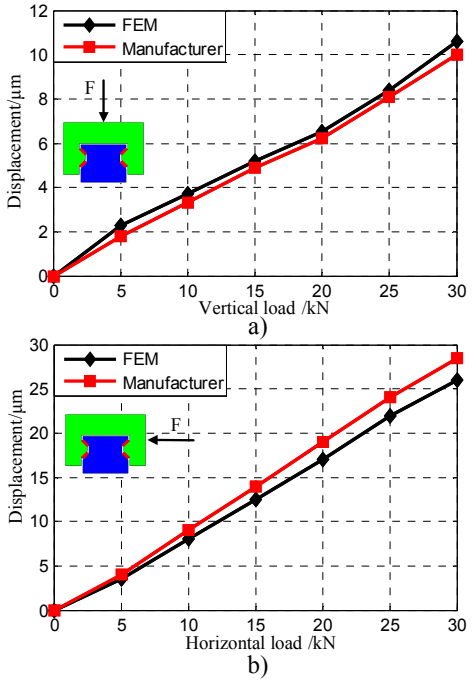

Fig.6. Comparison of slider stiffness curves.

As shown in Fig.6., the stiffness curves obtained from finite element analysis and provided by the guide rail manufacturers were compared [17]. It could be seen from the figure that the two curves basically coincided. The difference was in the range of 4.8 %~10.2 %. Therefore, it could be considered that the achieved qualitative and quantitative conformity were high. The correctness of the finite element model of guide rails was proved.

C. Integrated finite element model

The integrated finite element model consists of guide rails, slide blocks, and test plate. A fixed constraint was applied to the base of the rail. The material of the test plate was gray cast iron with Young´s modulus of 157 GPa, Poisson ratio of 0.27 and density of 7800 Kg/m^3. The tetrahedral element was used for meshing. The finite element model of test plate consisted of 3 765 elements and 7 265 nodes.

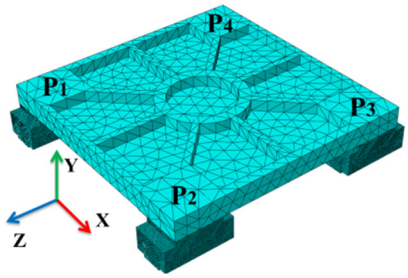

Fig.7. Integrated finite element model.

The bottom surface of the test plate and the upper surface of the slide block were tied together. Because the mesh size of the test plate was larger than slide blocks, the former was taken as a master surface and the latter was taken as a slave surface. The integrated finite element model is established in Fig.7., where P_1, P_2, P_3, P_4 are the serial numbers of the sliders. In the finite element analysis, the boundary condition of the guide rails was changed to make the slide blocks move in the X and Y directions so as to analyze the influence of geometric errors on the stress state of the test plate.

Due to the defects in the casting process, the material distribution was uneven. In the modeling process, the tiny characteristics of the test plate were neglected, which led to the deviation between the finite element model and the actual model. By comparing the measured strain and the strain obtained from the finite element analysis, Young´s modulus and density of the test plate were modified repeatedly in the finite element model. Finally, Young´s modulus and density were determined to be 126 GPa and 7450 Kg/m^3.

4. SELECTION OF THE KEY MEASUREMENT POINTS ON TEST PLATE

The stress state of the test plate is related to the relative displacements of slide blocks. The relative displacements of slide blocks equal to the variation of the guide rail geometric errors. As shown in Fig.8., the slide blocks move along the Z axis, with the fourth slide block as the benchmark.

Consequently, we can obtain the displacements Δx_1, Δy_1, Δx_2, Δy_2, Δx_3, Δy_3 of the P_1, P_2, P_3 slide blocks relative to the benchmark. Based on the finite element model, the stress state of the test plate was analyzed when the relative displacements of the slide blocks were changed.

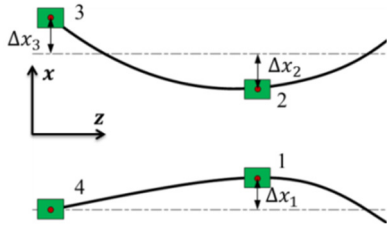

Fig.8. Relative displacements between slide blocks.

A. The key measurement points

With the change of the geometric errors, the relative position of slide blocks in the direction of X and Y will be changed, which leads to the deformation and the stress on the test plate [18]-[19]. But the stress was not evenly distributed. It was impossible to paste the strain gages at each location during the actual measurement process. Therefore, it was necessary to select some key measurement points. The stress at the points can adequately represent the stress state of the test plate and reflect the relative displacements of slide blocks.

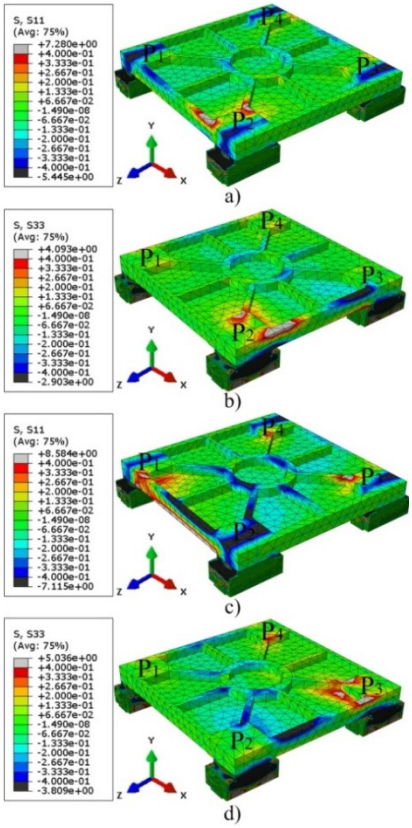

Fig.9. The stress nephogram of the test plate when the boundary conditions of the guide rails were changed. a) and c): the stress along the X direction, b) and d): the stress along the Z direction.

The stress nephograms in Fig.9.a) and Fig.9.b) were obtained by changing the boundary conditions of the rails in the finite element model, making Δy_2=5 μm, and the rest remained unchanged. Let Δx_2=5 μm and Δy_2=-5 μm to obtain the stress nephogram in Fig.9.c) and Fig.9.d).

Considering the stress state of the test plate, the positions where the stress changes more significantly were selected as the key measurement points as shown in Fig.10. The stress σ_x in the X direction was measured at positions 2, 4, 6, 8, 10, 12, 14, 16. Among them, 4, 8, 12 and 16 were located on the bottom of the test plate. The stress σ_z in the Z direction was measured at positions 1, 5, 9, 13. And the stress σ_{xz} along the direction of the 45° ribs was measured at positions 3, 7, 11 and 15.

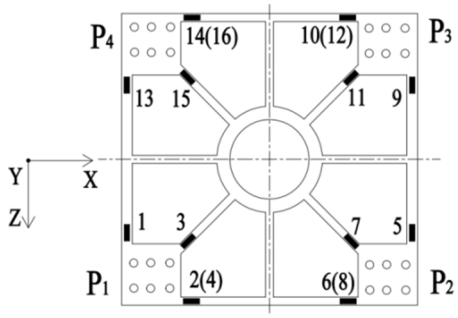

Fig.10. The key measurement points.

B. Resolution of measurement system

The measurement system consists of resistance strain gages, strain tester and static signal acquisition system. It was used to measure the strain at each key measurement point. From the knowledge of mechanics of materials, the relationship between stress σ and strain ε was:

$$\sigma = E\varepsilon \tag{3}$$

where E presents Young´s modulus of the test plate. Thus, the stress values at key measurement points can be obtained by (3).

Based on the integrated finite element model, a vertical downward displacement was applied to the guide rail at the position P_2 of the test plate. The displacement increases monotonically in the range of 1~10 μm, and the other positions of the guide rails did not impose displacement. The strain curves which were obtained by finite element analysis are plotted in Fig.11.

The results of finite element analysis show that only at points 2, 4, 5, 6, 7, 8 and 9 there is obvious strain. In other words, the strains at the other points were very small. As the test plate only has a bending strain, there is no tensile strain. The absolute value of the strain at points 4, 8 was equal to points 2, 6. Thus, it is not shown in Fig.11.

It can be seen from Fig.11. that 1 με was generated at the part of key measurement points on the test plate when the relative displacements of slide block were increased by 2 μm. The resolution of the resistance strain gage used in the measurement system was 1 με, while the strain tester had a resolution of 0.1 με. In other words, the resolution of the

measurement system was 2 μm. If a Wheatstone bridge or a higher resolution strain gage were used, the measurement system could identify a smaller relative displacement of slide blocks.

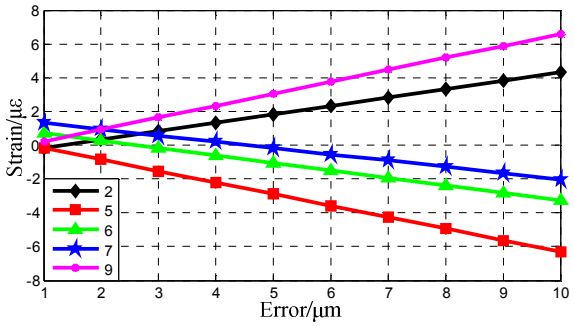

Fig.11. The change of strain curves.

5. BP NEURAL NETWORK

Nowadays, the neural network has become a more effective learning technique in pattern recognition areas. The neural network has a strong ability to learn and self-organize information. It only needs a few specific requirements and prior assumptions for modeling. These advantages have attracted much interest in the research of the machine error identification [20]-[21].

The stress state of the test plate is affected by the relative position of the four sliders. But it is difficult to calculate the stress of each point on the test plate by the relevant shell theory. Furthermore, there is a dynamic joint between the guide rail and the slider, and a static joint between the slider and the test plate too. Thus, it is difficult to establish the mathematical model between the stress and the geometric errors accurately. In this case, BP neural network is chosen to establish the mapping relationship between the stress and the guide rail geometric errors.

A. BP neural network

BP neural network does not need to know the structure and parameters of the object. Through the training of a number of learning samples, the mapping relationship between the input and output can be established. Fig.12. represents the structure of the neural network model including input layer, output layer and hidden layer.

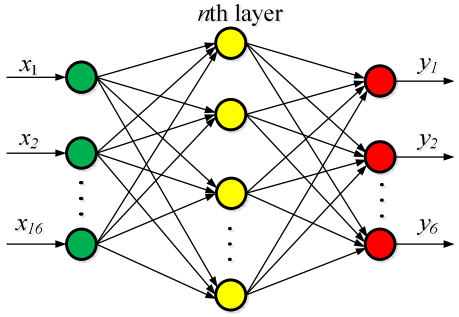

Fig.12. Structure of BP neural network.

Suppose there are m nodes in the input layer, l nodes in the output layer, and p nodes in the nth hidden layer. Taking the jth neuron of the nth layer as an example, the neuron has many inputs x_j^{n-1} which come from the neurons in the (n-1)th layer, but only a single output x_j^n that carries its signal to the neurons in the (n+1)th layer [22]. An adjustable weight, w_{ij}^n, represents the connecting strength between the ith neuron of the (n-1)th layer and the jth neuron of the nth layer. The output of the neural network is denoted as y_k (k=1, 2, ... l).

The hyperbolic tangent was selected as the transfer function of the neuron:

$$f(x) = \tanh(x) = \frac{1-\exp(x)}{1+\exp(x)} \tag{4}$$

Mathematically, the sum of the jth neuron net_j^n in the nth layer can be expressed as:

$$net_j^n = \sum_i w_{ij}^n x_{ij}^n - \theta_j^n \tag{5}$$

where θ_j^n was an internal threshold of the jth neuron. The output value x_j^n can be calculated as:

$$x_j^n = f(net_j^n) = \frac{1-\exp(-net_j^n)}{1+\exp(-net_j^n)} \tag{6}$$

From (4), (5) and (6), we can calculate the internal threshold θ_j and the adjustable weight w_{ij} of the neurons based on the input x_i. The training process is finished if the total error between the calculated output y_k and the desired output y_p is less than the given error value. Otherwise, θ_j and w_{ij} were adjusted according to the error back propagation algorithm.

The stress value $S_i(\sigma_1, \sigma_2, ... \sigma_{16})$ of the key measurement points in the finite element model was taken as the input and the slide block relative displacements $\Delta_i(\Delta x_1, \Delta y_1, \Delta x_2, \Delta y_2, \Delta x_3, \Delta y_3)$ were taken as the output. The BP neural network model was established using MATLAB as shown in Fig.12. The number of input and output nodes was 16 and 6, respectively. The number of hidden layers and nodes in each hidden layer were generally determined by a lot of experiments. After repeated experiments, the hidden layer was determined to be 2 layers and the number of nodes in the hidden layer was 12 and 8, respectively.

B. Training of BP neural network

Before the training of the BP neural network, we must first collect learning samples. The orthogonal test has the characteristics of neat comparability and equilibrium dispersion [23]. It can obtain more comprehensive samples as few as possible. Eight values of the displacement Δx_1, Δy_1, Δx_2, Δy_2, Δx_3 and Δy_3 were taken with respect to the reference sliders in the respective feasible domains. The orthogonal table L_{64} (8^9) was selected and there were 64 groups of learning samples. The learning samples were

obtained from the finite element model and the randomly selected 15 groups as the test samples. In order to accelerate the convergence of BP neural network, the learning samples should be normalized before training.

The maximum number of learning was 50,000, the learning rate was 0.01 and the permissible error was 10^{-6}. After the training of BP neural network, the test samples were used to test it. The results show that the difference is within 8 % as shown in Table 2. In this case, the mapping $f(\bullet)$ between S_i and Δ_i was established by $\Delta_i = f(S_i)$.

Table 2. Comparison of results.

	Output [μm]	Desired [μm]	Error [%]
Δx_1	5.3	5	6%
Δy_1	3.2	3	6.7%
Δx_2	-4.2	-4	5%
Δy_2	-5.4	-5	8%
Δx_3	8.5	8	6.3%
Δy_3	4.2	4	5%

The relative displacements of slide blocks $\Delta_i(\Delta x_1, \Delta y_1, \Delta x_2, \Delta y_2, \Delta x_3, \Delta y_3)$ can be obtained by inputting the stress values $S_i(\sigma_1, \sigma_2, \ldots \sigma_{16})$ into the trained BP neural network, and then getting a series of discrete points Δ_i. After the data processing, the geometric errors of two guide rails in the X and Y directions can be obtained.

6. Experimental verification

The BP neural network demonstrates the mapping relationship between the stress values $S_i(\sigma_1, \sigma_2, \ldots \sigma_{16})$ and the relative displacements $\Delta_i(\Delta x_1, \Delta y_1, \Delta x_2, \Delta y_2, \Delta x_3, \Delta y_3)$ of slide blocks. The measured stress values at the key measurement points of the test plate are input into the trained BP neural network. And the relative displacements of slide blocks which were obtained from the BP neural network are compared with the measured geometric errors. Thus, the correctness of the integrated finite element model and the practicability of BP neural network were verified.

A. Geometric error measurement of guide rails

As shown in Fig.13.a), the straightness of the two guide rails in the X and Y directions was measured by the photoelectric auto-collimator, respectively. The reflector was attached to the upper surface of the slide block. And the measuring light path between the reflector and the photoelectric auto-collimator was adjusted [24]. Move the slide block to the starting position and set the position as the reference point in the photoelectric auto-collimator. Move the slide block 100 mm at a time. The data was collected after the signal of the photoelectric auto-collimator stabilized. Thereby, we can obtain the straightness of each guide rail in the X and Y direction.

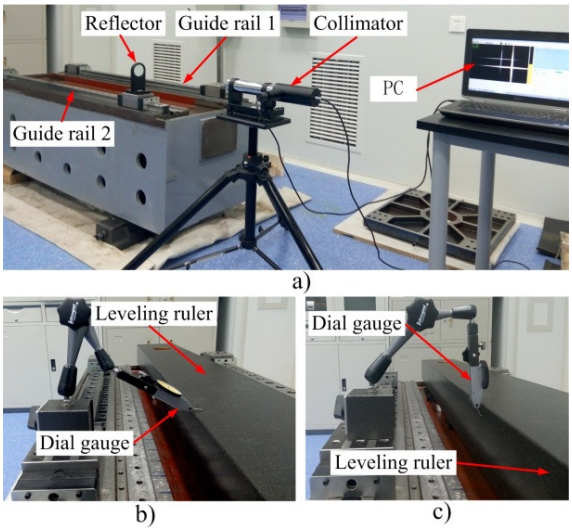

Fig.13. Guide rail geometric error measurement.

The high-precision leveling ruler and high-sensitivity dial gauge was used to measure the parallelism between the two guide rails. In Fig.13.b) and Fig.13.c), the leveling ruler was laid on the test bench and used as the measurement reference. The dial gauge was attached to the slide block, and the pointer was in contact with the top and side surface of the leveling ruler, respectively. Measurements were made from the starting position. The slide block was moved 100 mm at a time and then read the dial gauge. The above measurement procedure was repeated three times and the average value was taken. Taking the guide rail 1 as the benchmark, the geometric errors of the two guide rails in the X and Y direction can be obtained, as shown in Fig.14.

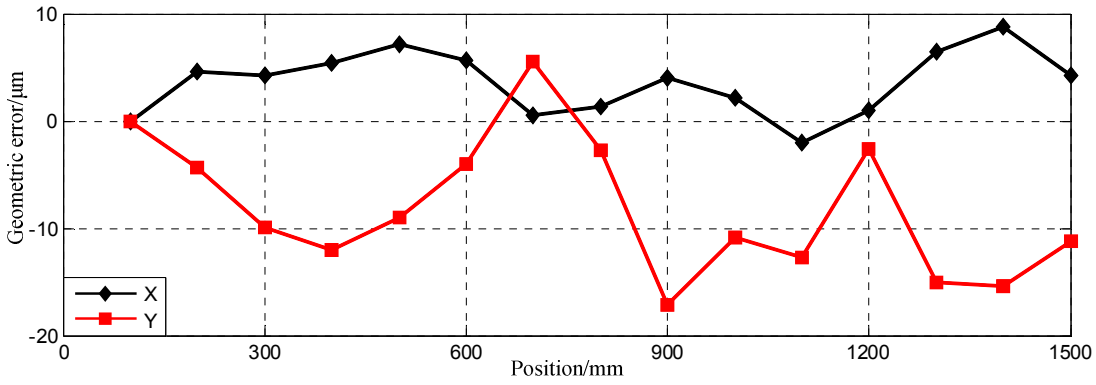

Fig.14. Geometric errors in the X and Y direction.

B. Stress measurement

The test plate was first bolted to the four slide blocks and a torque of 40 N•m was applied to the bolt with a torque wrench. Attach the strain gage to the surface of test plate and ensure that the orientation of the strain gage corresponds to the direction shown in Fig.10. Since the experiment was carried out in a constant temperature environment of 20±0.5 °C, no temperature compensation was required [19].

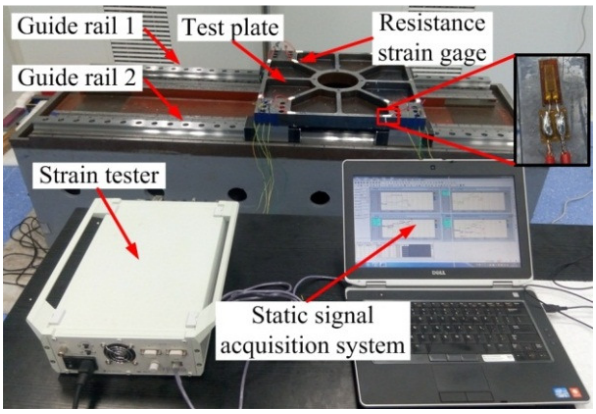

Fig.15. The stress measurement process.

The test plate was moved to the starting position. Then we balanced each channel and cleared the zero point. The test plate moved 100 mm every time. The data was collected after the signal was stable. By filtering out the interference signals during the measurement process, the strain relative to the starting position of the test plate can be obtained throughout the trip. The measurement process was repeated three times in order to obtain the average value of strain. The strain value was converted into stress according to (3).

Then we changed the tightening torque of the bolts. A tightening torque of 60 N•m was applied to the bolts with a torque wrench and the above measurement process was repeated. Comparing the stress values under different bolt tightening torques, it was found that the tightening torque of the bolts has little effect on the measurement results. In other words, bolt tightening torque has little influence on the relative displacement between slide blocks.

C. Comparison of the results

The stress values collected by the measurement system were input into the trained BP neural network to obtain the relative displacement of slide blocks. Taking the guide rail 1 as the benchmark, the relative displacements were converted into the geometric errors of the two guide rails in the X and Y directions.

In the X direction, the maximum error between the output value of BP neural network and the measured value was 4.2 μm. The maximum error value in the Y direction was 5 μm. Comparing the error curves in Fig.16.a) and Fig.16.b), we learned that the established BP neural network was practical. And the mapping relationship between the stress of the test plate and the geometric error of guide rail can be established by BP neural network.

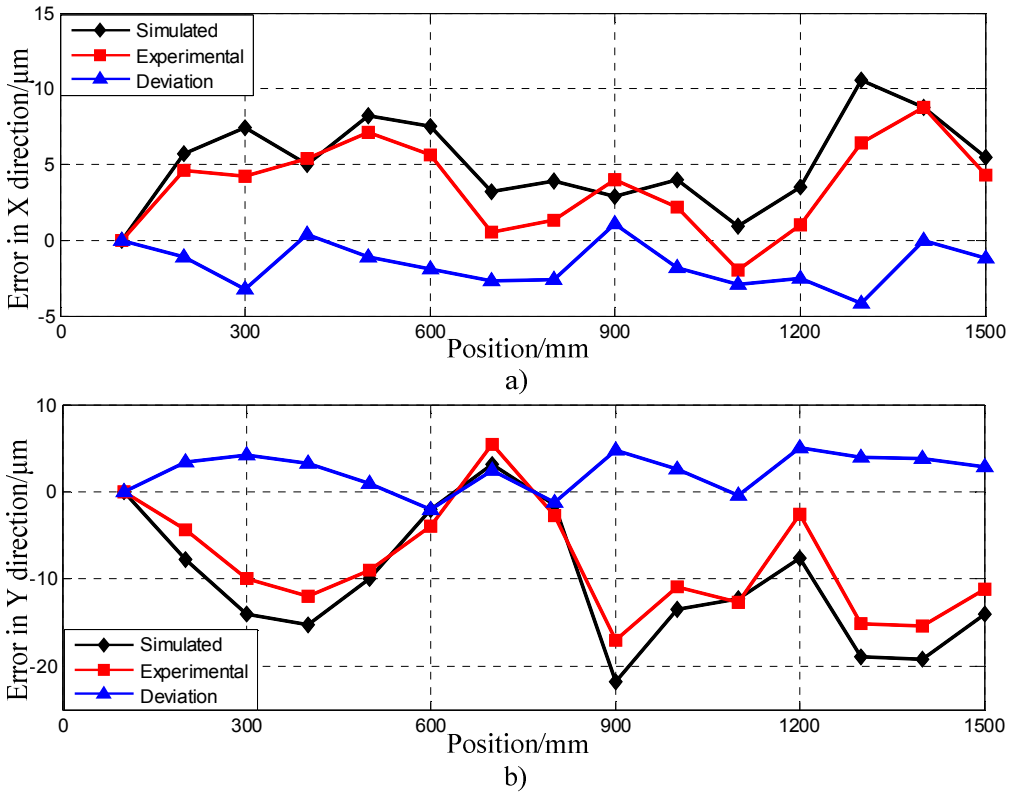

Fig.16. The output of the BP neural network and the measured geometric errors.

D. Geometric error adjustment

According to the results obtained from BP neural network, the geometric errors of the two guide rails in the X and Y directions were adjusted. The geometric errors in the X direction can be adjusted by adjusting the screw of the wedge plate or lapping the side mounting surface of the guide rail [6]. For the geometric errors in the Y direction, we generally use the scraping rail mounting surface to adjust the geometric errors.

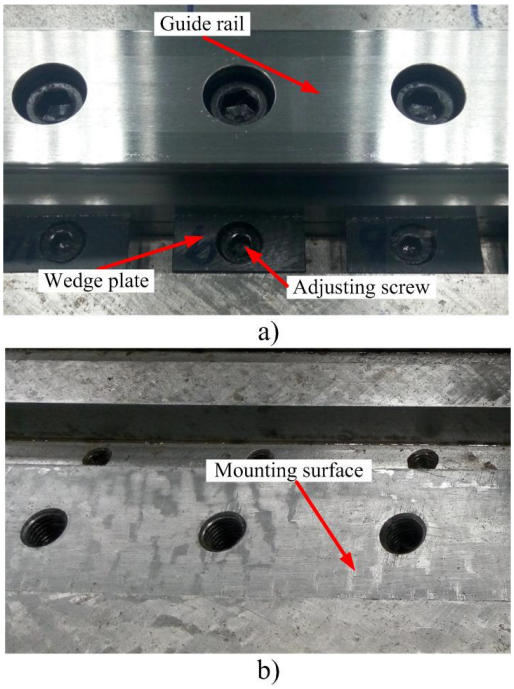

Fig.17. The adjustment of geometric errors.

When the geometric errors adjustment was completed, the stress at each key measuring point of the test plate was measured again. Then inputting the measured stress into the BP neural network, we can get the geometric errors in the X and Y direction again. According to the obtained error curves, the geometric errors of the guide rail were adjusted again until the stress did not exist at each key measurement point. At this time, the additional load of slide blocks during the movement was greatly reduced, which was helpful to improve the service life of the guide rail and the motion precision of the table.

7. CONCLUSION

A new method for guide rail geometric error identification and adjustment has been proposed. In finite element model, the sub-assemblies were discretized with solid elements and the roller was replaced by the spring element. Through the finite element analysis of the self-designed test plate, sixteen points on the surface of test plate were selected as the key measurement points, which the stress changed more significantly. The mapping relationship between the stress of the test plate and the geometric errors of the guide rail was established by using the BP neural network. The data extracted from the finite element model was used as the learning samples to train the BP neural network. The test error was no more than 10 %, indicating that the BP neural network has good generalization ability. Compared to the geometric errors that we obtained from BP neural network and the measurement, the maximum difference was 4.2 μm and 5 μm in the X and Y direction, respectively. In conclusion, the proposed method can accurately identify the geometric errors of the guide rail and provide reference for geometric errors adjustment.

ACKNOWLEDGMENT

This research project is supported by the National Natural Science Foundation of China (No.51675378) and the National Science and Technology Major Project of China (No.2015ZX04005001).

REFERENCES

[1] Okafor, A.C., Ertekin, Y.M. (2000). Derivation of machine tool error models and error compensation procedure for three axes vertical machining center using rigid body kinematics. *International Journal of Machine Tools & Manufacture*, 40 (8), 1199-1213.

[2] Tian, W., Gao, W., Zhang, D., Huang, T. (2014). A general approach for error modeling of machine tools. *International Journal of Machine Tools & Manufacture*, 79 (4), 17-23.

[3] Ohta, H., Tanaka, K. (2010). Vertical stiffnesses of preloaded linear guideway type ball bearings incorporating the flexibility of the carriage and rail. *Journal of Tribology*, 132 (1), 547-548.

[4] Rahmani, M., Bleicher, F. (2016). Experimental and numerical studies of the influence of geometric deviations in the performance of machine tools linear guides. *Procedia CIRP*, 41, 818-823.

[5] Shamoto, E., Park, C.H., Moriwaki, T. (2001). Analysis and improvement of motion accuracy of hydrostatic feed table. *CIRP Annals - Manufacturing Technology*, 50 (1), 285-290.

[6] Khim, G., Park, C.H., Shamoto, E., Kim, S.W. (2011). Prediction and compensation of motion accuracy in a linear motion bearing table. *Precision Engineering*, 35 (3), 393-399.

[7] Khim, G., Oh, J.S., Park, C.H. (2014). Analysis of 5-DOF motion errors influenced by the guide rails of an aerostatic linear motion stage. *International Journal of Precision Engineering and Manufacturing*, 15 (2), 283-290.

[8] Kim, G.H., Han, J.A., Lee, S.K. (2014). Motion error estimation of slide table on the consideration of guide parallelism and pad deflection. *International Journal of Precision Engineering and Manufacturing*, 15 (9), 1935-1946.

[9] Li, J., Mao, K., Chen, Q., Nie, Y. (2015). Experimental research of large components of machine tools assembly stress distribution under different assembly process. *Machine Tool & Hydraulics*, 43 (21), 118-122.

[10] Bosetti, P., Bruschi, S. (2012). Enhancing positioning accuracy of CNC machine tools by means of direct measurement of deformation. *The International Journal of Advanced Manufacturing Technology*, 58 (5), 651-662.

[11] Liu, Y., Liu, M., Yi, C., Chen, M. (2014). Measurement of the deformation field for machine tool based on optical fiber Bragg grating sensors. In *International Conference on Innovative Design and Manufacturing*, 13-15 August 2014. IEEE, Vol. 971-973, 222-226.

[12] Chlebus, E., Dybala, B. (1999). Modelling and calculation of properties of sliding guideways. *International Journal of Machine Tools & Manufacture*, 39 (12), 1823-1839.

[13] Majda, P. (2012). Relation between kinematic straightness errors and angular errors of machine tool. *Advances in Manufacturing Science & Technology*, 36, 47-53.

[14] Majda, P. (2012). Modeling of geometric errors of linear guideway and their influence on joint kinematic error in machine tools. *Precision Engineering*, 36 (3), 369-378.

[15] Shi, Y., Zhao, X., Zhang, H., Nie, Y., Zhang, D. (2016). A new top-down design method for the stiffness of precision machine tools. *The International Journal of Advanced Manufacturing Technology*, 83 (9), 1887-1904.

[16] Zhupanska, O.I. (2011). Contact problem for elastic spheres: Applicability of the Hertz theory to non-small contact areas. *International Journal of Engineering Science*, 49 (7), 576-588.

[17] THK Co., Ltd. (2008). *THK Linear Motion System Catalog*.

[18] Kowalik, M., Rucki, M., Paszta, P., Gołębski, R. (2016). Plastic deformations of measured object surface in contact with undeformable surface of measuring tool. *Measurement Science Review*, 16 (5), 254-259.

[19] Gawedzki, W., Tarnowski, J. (2015). Design and testing of the strain transducer for measuring deformations of pipelines operating in the mining-deformable ground environment. *Measurement Science Review*, 15 (5), 256-262.

[20] Fuh, K.H., Wang, S.B. (1997). Force modeling and forecasting in creep feed grinding using improved BP neural network. *International Journal of Machine Tools & Manufacture*, 37 (8), 1167-1178.

[21] Basheer, I.A., Hajmeer, M. (2000). Artificial neural networks: Fundamentals, computing, design, and application. *Journal of Microbiological Methods*, 43 (1), 3-31.

[22] Rafiq, M.Y., Bugmann, G., Easterbrook, D.J. (2001). Neural network design for engineering applications. *Computers & Structures*, 79 (17), 1541-1552.

[23] Tsui, K. (2007). Strategies for planning experiments using orthogonal arrays and confounding tables. *Quality & Reliability Engineering*, 4 (2), 113-122.

[24] Ekinci, T.O., Mayer, J.R.R. (2007). Relationships between straightness and angular kinematic errors in machines. *International Journal of Machine Tools & Manufacture*, 47 (12-13), 1997-2004.

Received February 02, 2017.
Accepted May 18, 2017.

MEASUREMENT SCIENCE REVIEW, **17**, (2017), No. 3, 145-152

DE GRUYTER OPEN

MEASUREMENT SCIENCE REVIEW

Journal homepage: http://www.degruyter.com/view/j/msr

ISSN 1335-8871

Magnetic Resonance Super-resolution Imaging Measurement with Dictionary-optimized Sparse Learning

Jun-Bao Li[1], Jing Liu[2], Jeng-Shyang Pan[3], Hongxun Yao[4]

[1]*Department of Automatic Test and Control, Harbin Institute of Technology, Harbin 150080, China*
[2]*College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China*
[3]*Fujian Provincial Key Lab of Big Data Mining and Applications, Fujian University of Technology, Fuzhou 350108, China*
[4]*School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China*

Magnetic Resonance Super-resolution Imaging Measurement (MRIM) is an effective way of measuring materials. MRIM has wide applications in physics, chemistry, biology, geology, medical and material science, especially in medical diagnosis. It is feasible to improve the resolution of MR imaging through increasing radiation intensity, but the high radiation intensity and the longtime of magnetic field harm the human body. Thus, in the practical applications the resolution of hardware imaging reaches the limitation of resolution. Software-based super-resolution technology is effective to improve the resolution of image. This work proposes a framework of dictionary-optimized sparse learning based MR super-resolution method. The framework is to solve the problem of sample selection for dictionary learning of sparse reconstruction. The textural complexity-based image quality representation is proposed to choose the optimal samples for dictionary learning. Comprehensive experiments show that the dictionary-optimized sparse learning improves the performance of sparse representation.

Keywords: Magnetic resonance imaging measurement, sparse learning, dictionary learning, super-resolution imaging.

1. INTRODUCTION

Magnetic Resonance Super-resolution Imaging Measurement (MRIM) is widely used in physics, chemistry, biology, geology, medical and material science, and especially in medical diagnosis. In the medical disease diagnosis, Magnetic Resonance Imaging (MRI) can image the organs and structures of body for clinical diagnosis. It performs better than X-ray, ultrasound, or computed tomography (CT) in the diagnosis of tumors, bleeding, or blood vessel diseases. MRI plays gradually a more important role in the diagnosis of various diseases. The current method is to improve the resolution through increasing free water magnetization of human tissue and organs. Accordingly, the method of increasing radiation time and radiation intensity of electromagnetic waves is widely applied to MR instruments. But the human body can be excessively heated by radiation which inactivates proteins. So, the hardware super-resolution imaging technology has its limitations on the practical clinical application. MRI can obtain the high-resolution image for the clinical diagnosis, but is limited by SNR, hardware, imaging time, and so on. Many methods are proposed to reconstruct the high-resolution image through signal processing and other machine learning methods. Sparse representation-based super-resolution reconstruction is a recently proposed method for image recovery. The sparse representation model-based image processing performs well on image denoising [1], image deblurring [2], [3], image restoration [4]. The sparse representation methods include rapid sparse representation [5], dictionary-based learning method, for example KSVD [6], MOD [7], pixel selection-based sparse representation [8], locality constrained sparse representation [9], precise dictionary representation [10], dictionary selection-based sparse representation [11]. Sparse representation methods are applied in many image processing techniques, including object detection via hyper spectral image [12], image fusion and restoration [13], and other image restoration [14], [15], image classification [16], and SR-based image classification and face recognition [17] techniques. In the previous works, image sparse super-resolution technologies were widely applied to medical image super-resolution, remote hyperspectral imaging, video and image super-resolution. The features of edge, texture, and structure are applied to image super-resolution [18]. For the dictionary training problem, the constraint dictionary method is proposed to SR-based image super-resolution [19], and the sparse domain based image deblurring is to solve the high-resolution image [20]. For SR-based medical image analysis, only a few SR-based medical image analysis methods were proposed in previous works, for example,

DOI: 10.1515/msr-2017-0018

sparse representation based MR spectroscopy quantification [21], constrained generative regression model-based fMRI analysis [21], filter-base machine intelligence [22]-[24], sparse coding based super-resolution learning [25], similar based image blocks sparse relations [26].

Based on the survey of the recent works on sparse reconstruction technologies, there are many researches on dictionary training, sparse parameters solution, sparse reconstruction. But to dictionary-optimized learning, less attention was paid in the previous works. They applied the traditional methods to training the dictionary with enough training samples. However, in many practical applications, i.e., MR, only limited samples are used to train the dictionary. MR images have to be used to train the dictionary for the excellent performances of super-resolution. How to sufficiently use the definite MR samples is a crucial issue of improving the MR super-resolution. The performance of dictionary learning training directly affects the quality of the image reconstruction. Thus, how to optimize the training procedure is a crucial problem. We have to select optimal samples to ensure the effectiveness of training dictionary blocks. In this paper, we present a framework of dictionary training method based on optimizing the samples from the limited MR training samples. Texture-complexity based optimization is to choose the training samples from the training samples. The texture complexity is measured with the gray-consistency method. Based on this dictionary training method, we propose a framework of dictionary training samples-optimized sparse reconstruction-based super-resolution MR imaging.

2. FRAMEWORK AND ALGORITHM

A. Framework

The framework is shown in Fig.1. In this framework, the coupled dictionaries are trained with sample selection via the image quality representation based on the gray-consistency method. The MR image training samples are selected for training the high-low MR image blocks from the training image samples, and these training samples are simultaneously trained low-high resolutions of dictionaries. The framework applies the machine learning-based dictionary block learning. In the definite scale of high resolution images, the optimized chosen high resolution images are spliced into multiple image blocks to train the high-resolution dictionary. Accordingly, the low-resolution dictionary is trained by the low image down-sampled high resolution image. In the same way, the multiple image blocks are achieved through splicing the low resolution training images. Under the SR-based super-resolution construction framework, the SR coefficients are solved with sparse representation constrained optimization equation, accordingly, the image can be represented by the combinations of the image block from the dictionary under the SR parameters. On the SR construction, the high resolution of image is computed under the sparse coefficients and the high-resolution dictionary.

As shown in the framework, the crucial issue is to choose the optimized training samples for dictionary training. For the SR construction-based MR super-resolution, dictionary training often depends on a large number of training samples, but in the practical applications we can obtain a sufficient number of training samples, in other words, among the definite number of training samples some training samples are not effective for dictionary training, so that the dictionary is not ideal for sparse representation of image. In the framework, the crucial step is the training sample selection, and the complexity-based image quality representation is presented for MR image sample selection for training the high-low MR image blocks. If the texture complexity of MR training images is higher, then the quality of super-resolution is better. Not all MR images are effective for the dictionary training. We apply the gray consistent-based complexity measuring method to discriminately classify the MR image sample. On the basis of discriminant MR training sample, the optimal training images are selected for the training dictionary, the samples that are not fitted to the dictionary training of MR images are deleted, which improve the training performance of dictionary.

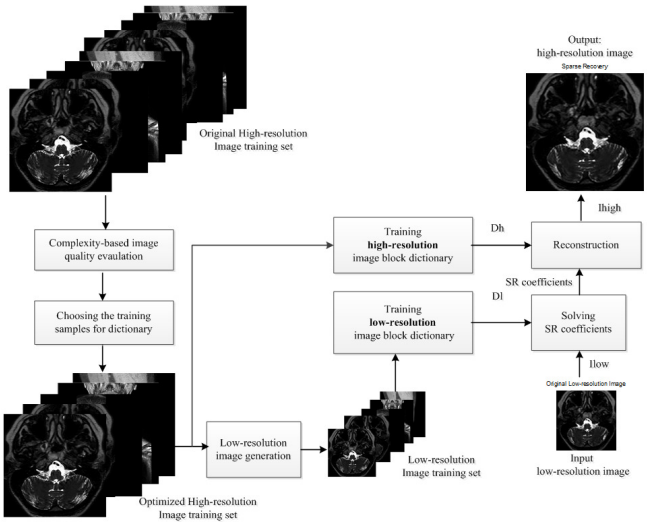

Fig.1. Framework of Dictionary Training Samples-Optimized Sparse Reconstruction-Based Super-resolution MR Imaging.

B. Algorithm

Firstly, we describe the algorithm of the sample selection for dictionary training. In this algorithm, the complexity of MR image depends on the type of object and representation method. The description of image complexity is based on the angle of the whole, the angle of the region and the angle of the target. So, the spatial distribution of gray level is unique to the image, and the two-dimensional image will not be related to the spatial location. The gray distribution reflects the spatial distribution of the image, which describes the size and the spatial distribution of the gray patches. The distribution of gray space is used to describe the image correlation and symmetry. The features include concentrated

or dispersed gray consistency, the existence of repetition, symmetry and so on. The consistency of gray level is used to describe the complexity of the image as follows

$$U = \sum_{a=2}^{m-1}\sum_{b=2}^{n-1}(f(a,b)-\bar{f})^2 \tag{1}$$

$$\bar{f} = \frac{1}{8}\left(\sum_{i=-1}^{1}\sum_{j=-1}^{1} f(a+i,b+j) - f(a,b)\right) \tag{2}$$

where $f(a,b)$ is the pixel value in the position, $\bar{f}$ is the mean value of 8 pixels round the center. Gray consistency shows the difference of each pixel and the pixel gray value accumulation. The greater is the difference in description of the image pixel gray value wind speed change, the more complex is the performance of the different image texture. We implement some experiments on the performance of the gray consistency method.

The algorithm procedure is shown in Table 1. If the image samples based on texture complexity, or most of the regional images as the invalid area that is adjacent to the pixel point gradient value are selected, the human eye cannot distinguish between the image patches. So randomly selected is invalid area. Based on texture fuzzy region, more samples improve the performance of super-resolution reconstruction. Therefore, not all training samples are necessary for dictionary training, and only the training image sample with the complex textures can provide enough features of the image blocks for the dictionary training for SR.

Table 1. Algorithm procedure of training images selection.

Step 1. Compute the largest gradient and complexity of the training sample.
For the input image I_i, compute the first order and second order $f(I_i)$:

$$\begin{cases} f_1 = [-1,0,1], & f_2 = f_1^T \\ f_3 = [1,0,-2,0,1], & f_4 = f_3^T \end{cases}$$

Compute the complexity of all images I_i:

$$y_i = \sum_{a=2}^{m-1}\sum_{b=2}^{n-1}(I_i(a,b)-\bar{I}_i)^2, \quad \bar{I}_i = \frac{1}{8}\left(\sum_{i=-1}^{1}\sum_{j=-1}^{1} I_i(a+i,b+j) - I_i(a,b)\right)$$

where m and n are the number of row and column, and $I_i(a,b)$ is the pixel gray value of the point (a,b).
Step 2. Select the center line of two coordinates as the baseline.
Step 3. Project the center of each class to the baseline, and cluster the image data to the center of each class.
Step 4. Select the sample of the farthest distance from the center as the training sample.

Secondly, the algorithm of dictionary-learning super-resolution method is described as follows. The high-resolution image block dictionary D_h and the low-resolution image block dictionary D_l are trained by high-resolution and low-resolution of MR images. As the definition of SR method, the high-resolution I_{high} contains many image blocks B_{high}, which is represented as a sparse linear combination of the D_h and sparse representation parameter vector α as

$$B_{high} \approx D_h\alpha \tag{3}$$

The α is to recover B_{low} from the low-resolution MR image I_{low} according to the low and high resolution of D_l and D_h with the following optimization equations as follows.

$$\min \|a\|_1 \quad s.t. \|FD_l a - FB_{low}\|_2^2 \le \varepsilon \tag{4}$$

where F denotes the feature extractor for dictionary generation, which provides the constraint of similarity of the coefficients a and B_{low}. F improves the high prediction accuracy through computing coefficients. The high-pass filter is to extract the features because of the sensitivity of human vision. The high-frequency components of low-resolution image are to predict the high-frequency parts of high-resolution of MR image. Then the optimization equation (5) is transferred to the Lagrange problem:

$$\min_a \|FD_l a - FB_{low}\|_2^2 + \lambda\|a\|_1 \tag{5}$$

where λ is balance sparsity of B_{low}., and λ determines the construction performance. The super-resolution reconstruction $D_h\alpha$ of B_{low} is calculated adjacent B_{high}, so

$$\min \|a\|_1 \quad s.t. \|FD_l a - FB_{low}\|_2^2 \le \varepsilon_1 and \|MD_l a - m\|_2^2 \le \varepsilon_2 \tag{6}$$

where M extracts the overlapped region between the target block and previously reconstructed MR image. Supposed

$\overline{D} = \begin{bmatrix} FD_l \\ \beta MD_h \end{bmatrix}$, $\overline{B}_{low} = \begin{bmatrix} FB_{low} \\ \beta m \end{bmatrix}$, then

$$\min_a \|\overline{D}a - \overline{B}_{low}\|_2^2 + \lambda\|a\|_1 \tag{7}$$

The high-resolution block B_{high} is reconstructed with the optimal solution α^* of optimization equation (5) as follows.

$$B_{high} = D_h a^* \tag{8}$$

For the noised image, $I_{orginal}$ may not satisfy the reconstruction constraint of sparse representation-based construction. Under the blurring operator F_B and down-sampling operator F_S, I_{low} is achieved by high-resolution I_{high} as follows.

$$I^* = \arg\min_I \|F_S F_B I - I_{low}\|_2^2 + c\|I - I_{original}\|_2^2 \tag{9}$$

We applied gradient descent method to solve the constraint optimization equation based on the iteration method as

$$I(n+1) = I(n) + \eta[F_B^T F_S^T (I_{low} - F_S F_B I(n)) + c(I(n) - I_{original})] \tag{10}$$

where $I(n)$ is nth iteration of high-resolution MR image under the gradient step η. Finally, the solution I^* is the high-resolution reconstructed image of MR image $I_{orginal}$. So, the high-resolution of MR image I^* is reconstructed by α^* under the representations-based reconstructed constrained equation as

$$I^* = \arg\min_{I,\{a_{ij}\}} \left\{ \left\| F_S F_B I - I_{low} \right\|_2^2 + \lambda \sum_{i,j} \left\| a_{ij} \right\|_0 + \gamma \sum_{i,j} \left\| D_h a_{ij} - P_{ij} I \right\|_2^2 + \tau \rho(I) \right\} \quad (11)$$

Under the sparse representation coefficients α_{ij} in the (i,j) patch of I, and a penalty function $\rho(I)$ of encoding additional prior knowledge on the high-resolution and low resolution MR image.

3. RESULTS

In this section, we implement some experiments to testify the feasibility of the algorithm, and have the comprehensive evaluation of the performance of the algorithm. We apply real medical MR images. Some examples are shown in Fig.2. The images include brain, ankle, aorta, carotid artery, knee, neck, and foot. The size of the training dictionary is 125×512, and 512 feature blocks, 5×5 of image blocks, and overlap block is 4, balance parameter λ=0.1, and super-resolution rate is 1:2 and 1:4. The PSNR is to measure the super-resolution performance. In the feasibility of the dictionary training, in order to analyze each dictionary, we randomly choose a few of the training samples for dictionary training, and 1, 2, 3, 5, 10, 20, 105, 360 of MR images are chosen randomly to construct the training dictionary sets. In the performance evaluation, we repeated the experiments with the Monte-Carlo simulation method.

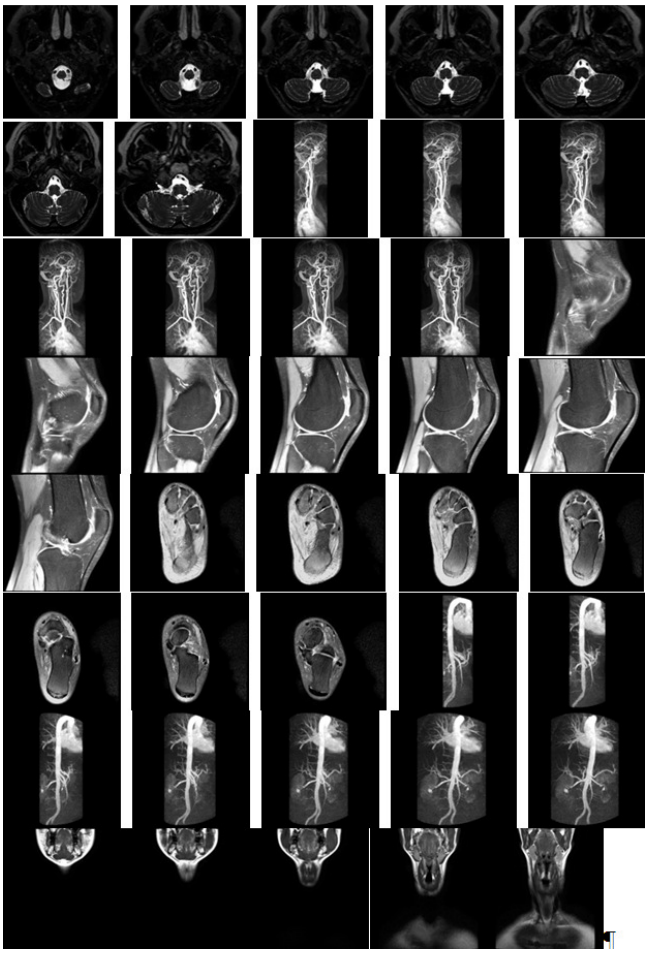

Fig.2. Examples of MRI sequence.

Table 2. Performance on MR super-resolution under the different training number.

	The number of training samples	**Dictionary**	**brain**	**ankle**	**aorta**	**carotid artery**	**knee**	**neck**	**foot**
1:2 of super-resolution	1	D1	31.77	28.71	35.70	31.72	30.51	30.19	30.42
	2	D2	32.15	28.80	35.97	31.97	30.58	30.25	30.74
	3	D3	32.31	28.84	36.05	32.07	30.60	30.31	30.87
	5	D4	32.25	28.85	36.05	32.02	30.59	30.33	30.78
	10	D5	32.28	28.88	36.05	32.04	30.57	30.33	30.79
	20	D6	32.27	28.87	36.02	32.05	30.59	30.33	30.79
	105	D7	32.31	28.87	36.05	32.04	30.58	30.36	30.81
	360	D8	32.21	28.86	35.98	32.00	30.56	30.29	30.71
1:4 of super-resolution	1	D9	29.39	25.20	32.71	29.00	27.08	27.46	27.88
	2	D10	30.05	25.41	32.96	29.23	27.11	27.54	28.23
	3	D11	29.63	25.33	32.84	29.13	27.10	27.53	28.19
	5	D12	29.83	25.39	32.90	29.20	27.09	27.56	28.20
	10	D13	29.92	25.44	32.99	29.28	27.12	27.62	28.28
	20	D14	29.86	25.38	32.94	29.26	27.11	27.59	28.27
	105	D15	29.84	25.40	32.97	29.25	27.10	27.61	28.22
	360	D16	29.95	25.42	32.98	29.34	27.14	27.67	28.26

A. On the dictionary optimization

In this section, we evaluate the dictionary generation based on optimizing the training samples for SR-based construction of MR image.

Secondly, 20 images from different classes are selected as training samples of dictionary generation, and a dictionary is trained by each image to generate 20 dictionaries for SR-based MR super-resolution image. In the experiments, the procedural parameters include that the size of training dictionary is 125×512, the characteristic of block is 512, the image block size is 5×5, the number of overlapped block is 4, the balance parameter is 0.1, and the rate of super-resolution is 1:2. Table 3. shows the performance of MR super-resolution under the different number of training samples for dictionary generation. The first column is the training image ID, the second is the dictionary ID, and the remaining columns are the PSNR performance under the different MR images. For each test image, the trend of PSNR changing corresponding to the number of training image is the same. The experiments show that the different training sample sets for dictionary generation affect the performance of SR-based construction, from the fact that the trend of PSNR corresponding to the 7 sets of test images is the same.

Thirdly, further analysis on the relationships between complexity of training images and reconstruction performance is shown in Table 4. The circle in the figure indicates that the PSNR is higher. The complexity of the dictionary is higher, better performance of image super-resolution reconstruction is achieved.

Table 3. Performance on the different number of training samples for dictionary learning (1:4 of MR super-resolution).

Image ID	Generated dictionary	brain	ankle	aorta	carotid artery	knee	neck	foot
1	D30	30.31	27.86	34.68	30.85	30.19	29.61	29.49
2	D31	31.95	28.57	35.69	31.78	30.40	29.98	30.49
3	D32	32.43	28.93	36.17	32.15	30.63	30.44	31.05
4	D33	32.14	28.79	35.98	31.93	30.60	30.37	30.81
5	D34	31.88	28.53	35.58	31.69	30.38	29.92	30.46
6	D35	32.23	28.83	36.04	32.02	30.61	30.31	30.79
7	D36	32.39	28.74	36.03	31.95	30.56	30.31	30.84
8	D37	30.47	27.94	34.80	30.98	30.21	29.69	29.74
9	D38	32.12	28.64	35.77	31.78	30.40	29.98	30.61
10	D39	32.44	28.96	36.17	32.09	30.65	30.46	30.96
11	D40	32.31	28.96	36.16	32.19	30.63	30.49	30.96
12	D41	31.88	28.54	35.59	31.68	30.35	29.91	30.49
13	D42	30.53	28.04	34.79	30.99	30.20	29.68	29.70
14	D43	30.06	27.65	34.63	30.79	30.18	29.58	29.55
15	D44	30.41	28.02	34.74	30.91	30.17	29.66	29.57
16	D45	32.20	28.87	35.96	31.94	30.60	30.33	30.75
17	D46	32.29	28.91	36.20	32.18	30.59	30.51	30.82
18	D47	31.94	28.54	35.66	31.70	30.39	29.93	30.48
19	D48	32.12	28.55	35.73	31.75	30.41	29.93	30.54
20	D49	32.01	28.64	35.80	31.76	30.42	30.07	30.53

Table 4. Relationship between the complexity and reconstruction performances.

ID	Dictionary	PSNR	Complexity (10^9)	ID	Dictionary	PSNR	Complexity (10^9)
1	D30	30.31	0.18	11	D40	32.31	0.89
2	D31	31.95	3.10	12	D41	31.88	1.18
3	D32	32.43	2.12	13	D42	30.53	0.19
4	D33	32.14	0.47	14	D43	30.06	0.25
5	D34	31.88	0.53	15	D44	30.41	0.19
6	D35	32.23	3.23	16	D45	32.20	1.54
7	D36	32.39	0.33	17	D46	32.29	1.60
8	D37	30.47	0.24	18	D47	31.94	2.09
9	D38	32.12	2.16	19	D48	32.12	1.44
10	D39	32.44	1.72	20	D49	32.01	1.24

B. On performance evaluation

We evaluate the proposed dictionary training samples-optimized sparse reconstruction-based MR images super-resolution. We repeated the experiments with the Monte-Carlo simulation method. In the experiments, we implemented 1:2 and 1:4 of image super-resolution MR to evaluate the performance of the MR sequence. The MR images are obtained through the down-sampling for performance evaluation. In the experimental procedural parameters, Overlaps is 4, λ=0.1. For the comparison, we implement three MR super-resolution methods including the traditional SR-based MR super-resolution, Bicubic Interpolation (BI), and the proposed optimal selecting training samples-based sparse representation construction of MR image. The performance of the super-resolution methods is measured by PSNR. The results are shown in Table 5. These results show that the proposed method is feasible to improve SR-based MR super-resolution. The results show that sufficient use of the definite MR samples is a crucial issue of improving the MR super-resolution. For joint dictionary learning, training samples will directly affect the quality of the image reconstruction effect of sparse image reconstruction, so the sample selection is one of the most important research problems in dictionary learning.

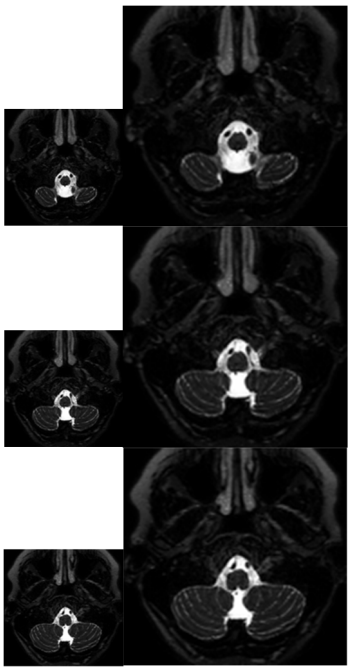

Fig.3. Examples of MR sequence super-resolution.

The traditional training method requires a large number of training samples to ensure the effectiveness of training dictionary blocks, more training samples possess relatively more prior knowledge. The trained dictionary can make the reconstruction of the image closer to the actual, but the sample is rich to ensure effective training samples. The poor quality of the samples not only provides high quality image reconstruction, it may also lower the value of the PSNR super resolution reconstruction. The proposed textural complexity-based image quality representation is feasible to select the optimal training images. The simultaneous training of low-high resolutions dictionaries is to enhance the similarity of sparse representations for sparse reconstruction-based super-resolution MR imaging. Some 1:2 of super-resolution examples are shown in Fig.3.

Other experiments are implemented on the real MR database from Harbin Medical University (HMU). This MR database includes 150 MR images, and we repeated the experiments with the Monte-Carlo simulation method. The PSNR value is to measure the performance of super-resolution. For the comparison, other super-resolution methods, Sparse Representation Reconstruction (SRR), Bicubic Interpolation (BI), and Neighborhood Embedding (NE), are used to compare the performance of the algorithms. The 1:2 and 1:4 of super-resolution rates are selected in the experiments. As results show in Table 6., the proposed algorithm achieves the highest performance. Learning-based super-resolution method is feasible to enhance the performance of image super-resolution.

Table 5. Performance evaluation of 1:2 and 1:4 of MR super-resolution.

	Methods	**1:4 super-resolution**	**1:2 super-resolution**
Skull Base	BI	32.54±1.54	34.65±1.26
	SRR	33.65±1.65	35.56±1.32
	Proposed	33.89±1.56	35.90±1.25
Carotid Arteries	BI	28.65±1.46	30.46 ±1.53
	SRR	29.56±1.56	31.65 ±1.68
	Proposed	30.25±1.44	32.25 ±1.56
Knee	BI	27.48±1.85	30.24±1.88
	SRR	28.56±1.65	31.36±1.86
	Proposed	29.86±1.75	32.56±1.36
Ankle	BI	27.36±1.66	30.52±1.83
	SRR	28.16±1.57	31.56±1.74
	Proposed	28.86±1.59	32.23±1.53
Ankle_pd	BI	27.42±1.58	30.25±1.85
	SRR	28.12±1.45	31.85±1.68
	Proposed	27.42±1.58	32.13±1.62
Aorta	BI	33.85±1.64	35.26±1.53
	SRR	34.12±1.56	36.46±1.56
	Proposed	34.97±1.34	36.96±1.63
Neck	BI	30.89±1.76	32.36±1.65
	SRR	31.35±1.53	33.12±1.76
	Proposed	32.23±1.86	33.86±1.86

Table 6. Performance evaluation on Harbin Medical University.

Methods	**1:2 super-resolution**	**1:4 super-resolution**
BI	30.26±0.76	27.52±0.64
NE	31.15±0.53	28.45±0.51
SRR	32.53±0.56	29.36±0.52
Proposed	32.86±0.65	29.78±0.56

4. CONCLUSION

We propose a framework of dictionary block-optimized sparse reconstruction-based super-resolution MR imaging. The paper aims to overcome the problem that the resolution of MR hardware imaging reaches the limitation of resolution due to the increasing of radiation intensity and time of magnetic exposure. The image quality representation based on complex procedures is presented for training the high-low MR image blocks. Comprehensive evaluations are implemented to test the feasibility and performance of the SR-MR method on the real database. Texture complexity of MR training images is higher, and then the quality of super-resolution is better. The optimal training images are selected for the training dictionary, the samples that are not fitted to the dictionary training of MR images are deleted. The super-resolution performance is improved. The proposed method can be applied to images super-resolution, video super-resolution applications. In the experiments, we use MR images of healthy volunteers to train the dictionary. But the coverage of all potential pathologies is not testified. The proposed algorithm improves the visual quality of MR images, but the performance of recovering the hidden pathology from the lower resolution of image is not testified. The feasibility on recovering the hidden information will be studied in the future work.

ACKNOWLEDGMENT

This work is supported by National Science Foundation of China under Grant No. 61671170, 61371178, Program for New Century Excellent Talents in University under Grant No. NCET-13-0168, Natural Science Foundation of Heilongjiang under Grant No. F2015003, Harbin Science and Technology Innovation Talent Research Special Fund (Youth Talent) under Grant No. 2015RQQXJ088, Guangxi Key Laboratory of Automatic Detecting Technology and Instruments under Grant No. YQ16201, the Open Projects Program of National Laboratory of Pattern Recognition, Fundamental Research Funds for the Central Universities under Grant No. HIT.BRETIII.201206, and Program for Interdisciplinary Basic Research of Science-Engineering-Medicine at the Harbin Institute of Technology.

REFERENCES

[1] Malathi, G., Shanthi, V. (2011). Statistical measurement of ultrasound placenta images complicated by gestational diabetes mellitus using segmentation approach. *Journal of Information Hiding and Multimedia Signal Processing*, 2 (4), 332-343.

[2] Lee, C.F., Chang, W.T. (2010). Recovery of color images by composed associative mining and edge detection. *Journal of Information Hiding and Multimedia Signal Processing*, 1 (4), 310-324.

[3] Kaganami, H.G., Ali, S.K. Zou, B. (2011). Optimal approach for texture analysis and classification based on wavelet transform and neural network. *Journal of Information Hiding and Multimedia Signal Processing*, 2 (1), 33-40.

[4] Hu, W.C., Yang, C.Y., Huang, D.Y., Huang, C.H. (2011). Feature-based face detection against skincolor like backgrounds with varying illumination. *Journal of Information Hiding and Multimedia Signal Processing*, 2 (2), 123-132.

[5] Peng, C.Y., Li, J.W. (2012). Fast sparse representation model for l1-norm minimisation problem. *Electronics Letters*, 48 (3), 162-164.

[6] Chavez-Roman, H., Ponomaryov, V. (2014). Super resolution image generation using wavelet domain interpolation with edge extraction via a sparse representation. *IEEE Geoscience & Remote Sensing Letters*, 11 (10), 1777-1781.

[7] Engan, K., Aase, S.O., Husey, J.H. (1999). Method of optimal directions for frame design. In *IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 5, 2443-2446.

[8] Li, Y., Namburi, P., Yu, Z., Guan, C. (2009). Voxel selection in fMRI data analysis based on sparse representation. *IEEE Transactions on Biomedical Engineering*, 56 (10), 2439-2452.

[9] Xu, D., Huang, Y., Zeng, Z., Xu, X. (2012). Human gait recognition using patch distribution feature and locality-constrained group sparse representation. *IEEE Transactions on Image Processing*, 21 (1), 316-326.

[10] Rubinstein, R., Bruckstein, A.M., Elad, M. (2010). Dictionaries for sparse representation modeling. *Proceedings of the IEEE*, 98 (6), 1045-1057.

[11] Cevher, V., Krause, A. (2011). Greedy dictionary selection for sparse representation. *IEEE Journal of Selected Topics in Signal Processing*, 5 (5), 979-988.

[12] Chen, Y., Nasrabadi, N.M., Trac, D. (2011). Sparse representation for target detection in hyperspectral imagery. *IEEE Journal of Selected Topics in Signal Processing*, 5 (3), 629-640.

[13] Yang, B., Li, S. (2010). Multifocus image fusion and restoration with sparse representation. *IEEE Transactions on Instrumentation and Measurement*, 59 (4), 884-892.

[14] Ogawa, T., Haseyama, M. (2011). Missing image data reconstruction based on adaptive inverse projection via sparse representation. *IEEE Transactions on Multimedia*, 13 (5), 974-992.

[15] Zuo, W., Lin, Z. (2011). A generalized accelerated proximal gradient approach for total-variation-based image restoration. *IEEE Transaction on Image Processing*, 20 (10), 2748-2759.

[16] Kang, L. W., Hsu, C.Y., Chen, H.W., Lu, C.S., Lin, C.Y., Pei, S.C. (2011). Feature-based sparse representation for image similarity assessment. *IEEE Transactions on Multimedia*, 13 (5), 1019-1030.

[17] Gao, S., Tsang, I.W., Chia, L.T. (2010). Kernel sparse representation for image classication and face recognition. *Lecture Notes in Computer Science*, 63 (14), 1-14.

[18] Yang, J., Wright, J., Huang, T. (2010). Image super-resolution via sparse representation. *IEEE Transactions on Image Processing*, 19, 2861-2873.

[19] Zeyde, R., Elad, M., Protter, M. (2012). On single image scale-up using sparse-representations. *Lecture Notes in Computer Science*, 6920, 711-730.

[20] Chavez-Roman, H., Ponomaryov, V. (2014). Super resolution image generation using wavelet domain interpolation with edge extraction via a sparse representation. *IEEE Geoscience & Remote Sensing Letters*, 11 (10), 1777-1781.

[21] Guo, Y., Ruan, S., Landre, J., Constans, J.M. (2011). A sparse representation method for magnetic resonance spectroscopy quantification. *IEEE Transactions on Biomedical Engineering*, 57 (7), 1620-1627.

[22] Wei, Y., Qiu, J., Karimi, H.R., Wang, M. (2014). H-infinity model reduction for continuous-time Markovian jump systems with incomplete statistics of mode information. *International Journal of Systems Science*, 45 (7), 1496-1507.

[23] Wei, Y., Qiu, J., Karimi, H.R., Wang, M. (2014). New results on H-infinity dynamic output feedback control for Markovian jump systems with time-varying delay and defective mode information. *Optimal Control, Applications and Methods*, 35 (6), 656-675.

[24] Wei, Y., Qiu, J., Karimi, H.R., Wang, M. (2015). Quantized H-infinity filtering for continuous-time Markovian jump systems with deficient mode information. *Asian Journal of Control*, 17 (6), 1914-1923.

[25] Yang, J., Wright, J., Huang, T. (2010). Image Super-resolution via sparse representation. *IEEE Transactions on Image Processing*, 19 (11), 2861-2873.

[26] Wang, J., Zhu, S., Gong, Y. (2010). Resolution enhancement based on learning the sparse association of image paths. *Pattern Recognition Letters*, 31 (1), 1-10.

Received March 29, 2017.
Accepted May 26, 2017.