USE OF INVERTIBLE RAPID TRANSFORM IN MOTION ANALYSIS

Ján GAMEC - Ján TURÁN
Dept. of Radioelectronics
Technical University of Košice
Park Komenského 13, 041 20 Košice
Slovakia

Abstract

The paper presents the results of a study on the use of invertible rapid transform for the motion estimation in a sequence of images. Novel motion estimation algorithms based on the analysis of the matrix of states (or matrices of states) of the invertible rapid transform are described. Experimental results with proposed algorithms are presented and compared with the well-known 2D-log method in real image sequences.

Keywords

motion estimation, rapid transform, invertible rapid transform

1. Introduction

The rapid transform (RT) [1] is a fast shift invariant transform. The RT is useful for pattern recognition, if the position of the pattern is unknown or the pattern is moving [2,3,4].

However, the RT eliminates non only knowledge about position but also a lot of information about the original pattern itself. Generally it is not possible to obtain the original pattern, only from its RT [1]. In [5, 6, 7] was introduced that exists an invertible fast shift invariant transform based on the RT. This transform (invertible rapid transform (IRT)) consists of an RT which supplies a shift invariant pattern from the input pattern and a binary coding process (generating additional data) which records the "phase-information" of the input pattern. Thus additional data are known as a matrix of states (binary matrix) for 1D-IRT or system of matrices of states (system of binary matrices) for 2D-IRT [5].

The proposed motion estimation algorithms [8, 9, 10, 11] are based on the presumption that in matrix of states or in system of matrices of states is included relevant information about an image. Cyclical translations in the image are deterministicaly and unambiguously encoded to the values of this matrix or system of matrices. These matrices may be computed with use of simple and very fast algorithm using operation of comparison, addition and subtraction [5].

At the proposed paper the methods and results of motion estimation based on using matrix of states or system of matrices of states will be described. The experimental results of these methods are compared with the results obtained by the well-known 2D-log method in real image sequences.

2. Rapid transform (RT)

RT is a member of the class CT (certain transforms - fast translation invariant transforms) [12]. The transforms of the class CT can be divided according to the employed functions $f_s(a,b)$, (s=1,2). In table 1 are transforms with various functions $f_s(a,b)$ [4, 12].

Tab. 1

	RT	NT	MT	QT
$f_1(a,b)$	a+b	max{a,b}	$a \vee b$	$(a+b)^2$
$f_2(a,b)$	|a-b|	min{a,b}	$a \wedge b$	$(a-b)^2$

The RT was introduced by Reitboeck and Brody [1] for application in pattern recognition.

The RT results from a simple modification of the Walsh-Hadamard transform (WHT). The signal flow graph for the RT is identical to that of WHT (Fig.1), except that the absolute value of the output of each stage of the iteration is taken before feeding it to the next stage.

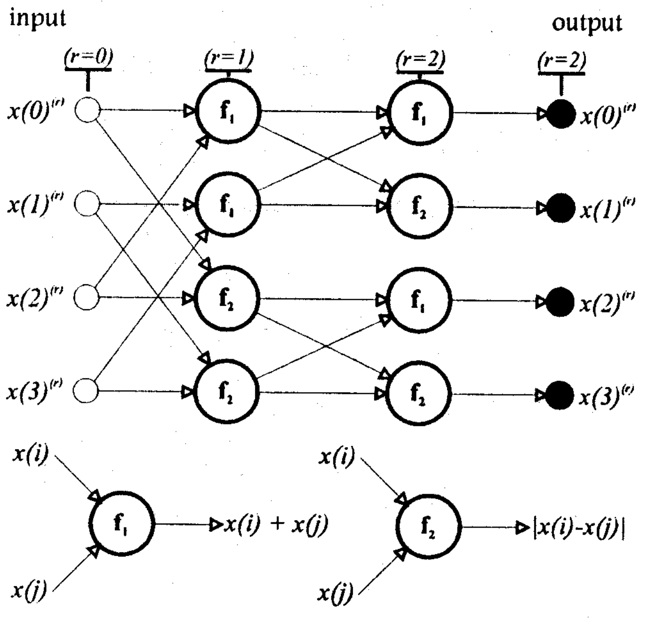


Fig.1 Signal flow graph of the 1D RT

This is not an orthogonal transform, as no inverse exists. RT has some interesting properties such as invariance to cyclic shift, reflection of the data sequence, and slight rotation of two-dimensional patterns. Various properties of RT have been developed in [2]. RT was used in recognition of alphanumeric characters [1, 2, 13, 14], robotics [4] and scene analysis [15]. More recently the modified rapid transform (MRT) [16] was presented to break undesired invariance of the RT, which leads to a loss of information about the original pattern. MRT was used in pattern recognition [16, 17] and in recognition of 3D objects [18]. In the following we will quickly review the RT.

Let us consider samples of one-dimensional signal represented as vector

$$\overline{\mathbf{X}} = \{x(i)\};\ i = 0,1,\dots,N\text{-}1 \quad N = 2n \tag{1}$$

and samples of two-dimensional signal represented as matrix

$$[\mathbf{X}] = \{x(i,j)\};\ i,j = 0,1,\dots,N\text{-}1 \quad N = 2^n. \tag{2}$$

In general transform data in step **r** are computed from data from step **r-1** with use of recurrent formula.

For one-dimensional signal

$$\begin{aligned} x(i+2js)^{(r)} &= \left| x(i+2js)^{(r-1)} + x[i+(2j+1)s]^{(r-1)} \right| \\ x[i+(2j+1)s]^{(r)} &= \left| x(i+2js)^{(r-1)} - x[i+(2j+1)s]^{(r-1)} \right| \end{aligned} \tag{3}$$

where $i = 0,1,\dots,s\text{-}1$,
$j = 0,1,\dots,t\text{-}1$,
$r = 1,2,\dots,n$,
$s = 2^{n-r}$, $t = 2^{r-1}$
n is the number of needed transform steps.

For two-dimensional signal

$$\begin{aligned} x(2i,2j)^{(r)} &= \Big\| x(i,j)^{(r-1)} + x(i+N/2,j)^{(r-1)} \Big| + \\ &\quad + \Big| x(i,j+N/2)^{(r-1)} + x(i+N/2,j+N/2)^{(r-1)} \Big\| \\ x(2i+1,j)^{(r)} &= \Big\| x(i,j)^{(r-1)} - x(i+N/2,j)^{(r-1)} \Big| + \\ &\quad + \Big| x(i,j+N/2)^{(r-1)} - x(i+N/2,j+N/2)^{(r-1)} \Big\| \\ x(2i,2j+1)^{(r)} &= \Big\| x(i,j)^{(r-1)} + x(i+N/2,j)^{(r-1)} \Big| - \\ &\quad - \Big| x(i,j+N/2)^{(r-1)} + x(i+N/2,j+N/2)^{(r-1)} \Big\| \\ x(2i+1,2j+1)^{(r)} &= \Big\| x(i,j)^{(r-1)} - x(i+N/2,j)^{(r-1)} \Big| - \\ &\quad - \Big| x(i,j+N/2)^{(r-1)} - x(i+N/2,j+N/2)^{(r-1)} \Big\| \end{aligned} \tag{4}$$

where $i, j = 0,1,\dots,(N/2\text{-}1)$,
$r = 0,1,\dots,n$
n is the number of needed transform steps.

3. Invertible rapid transform (IRT)

Even if RT is non-linear and thus non-invertible, by generating an additional data that indicates which pixel of an operand is greater during the forward RT, one can perform uniquely the invertible RT (IRT) [5, 6]. Thus additional data are known as a matrix of states $\boldsymbol{K}$ (binary matrix) for 1D-RT or system of matrices of states $\boldsymbol{K}_p^{(r)}$ (system of binary matrices) for 2D-RT. Signal [5] flow graph for compute of the 1D IRT is shown in Fig.2.

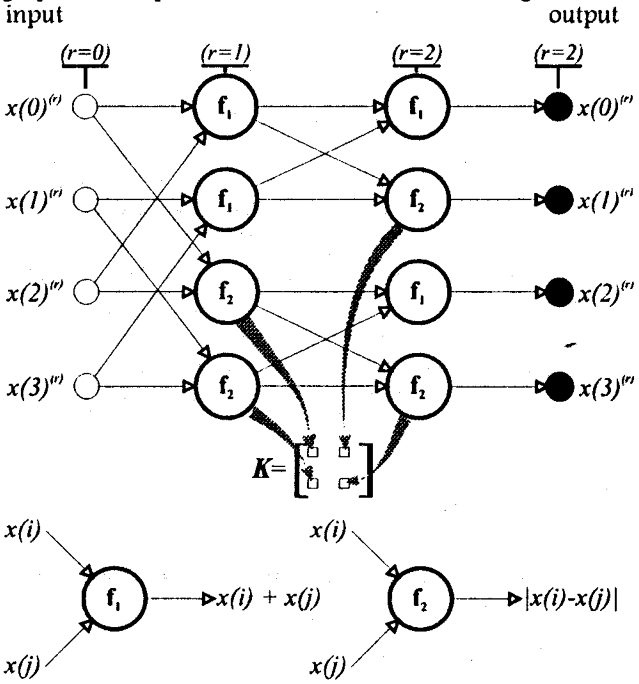


Fig.2 Signal flow graph of the 1D IRT

The matrix of states $\boldsymbol{K}$ or system matrices of states $\boldsymbol{K}_p^{(r)}$ may be computed as follows. For one dimensional case:

$$\begin{aligned} \boldsymbol{k}(i,r) &= 0, \text{ if } x(i)^{(r)} - x(i+N/2)^{(r)} < 0 \\ \boldsymbol{k}(i,r) &= 1, \text{ if } x(i)^{(r)} - x(i+N/2)^{(r)} \geq 0. \end{aligned} \tag{5}$$

The dimension of matrix $\boldsymbol{K}$ is $n \times N/2$. For two dimensional case:

$$k_1^{(r)}(i,j) = 1, \text{ if } x^{(r)}(i,j) - x^{(r)}(i+N/2,j) \geq 0$$

$$k_1^{(r)}(i,j) = 0, \text{ if } x^{(r)}(i,j) - x^{(r)}(i+N/2,j) < 0$$

$$k_2^{(r)}(i,j) = 1, \text{ if } x^{(r)}(i,j+N/2) - x^{(r)}(i+N/2,j+N/2) \geq 0$$

$$k_2^{(r)}(i,j) = 0, \text{ if } x^{(r)}(i,j+N/2) - x^{(r)}(i+N/2,j+N/2) < 0$$

$$k_3^{(r)}(i,j)=1,\text{ if } x^{(r)}(i,j)-x^{(r)}(i+N/2,j)\geq 0$$

$$k_3^{(r)}(i,j)=0,\text{ if } x^{(r)}(i,j)-x^{(r)}(i+N/2,j)<0$$

$$k_4^{(r)}(i,j)=1,\text{ if } x^{(r)}(i,j+N/2)-x^{(r)}(i+N/2,j+N/2)\geq 0$$

$$k_4^{(r)}(i,j)=0,\text{ if } x^{(r)}(i,j+N/2)-x^{(r)}(i+N/2,j+N/2)<0 \quad (6)$$

where *(r)* is transform step of IRT and

$i,j = 0,1,\dots,(N/2-1)$

$r = 1,2,\dots,n$

$p = 1,2,3,4$

The system of matrices of states $K_p^{(r)}$ is illustrated in Fig.3.

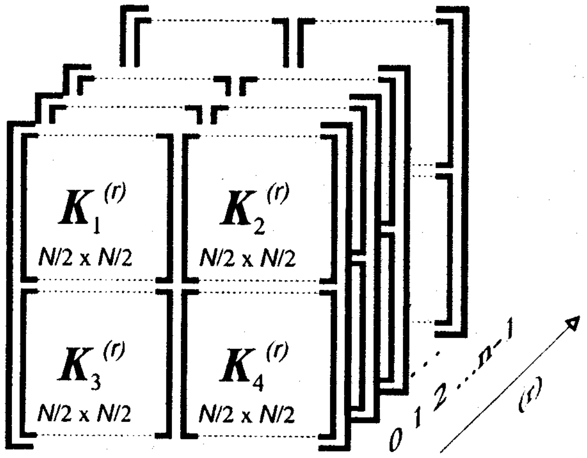


Fig.3 The system of matrices of states $K_p^{(r)}$

4. Motion estimation algorithms with use of IRT

This motion estimation algorithms are based on presumption that in matrix K or in system $K_p^{(r)}$ is included relevant information about the picture [2, 3, 4, 5, 9] and the motion in picture influence the first column of K or the first set of matrices of $K_p^{(r)}$ (i.e. $K_1^{(0)}, K_2^{(0)} K_3^{(0)}, K_4^{(0)}$) in maximal way.

First, the image is divided into smaller rectangular areas, which we call subblocks (See Fig.4). Let U_q be an N x N size subblock of frame **q** and U_{q-1} be equivalent subblock of frame **q-1**. Let search area (SA) be an $(N+2d_m)$ x $(N+2d_m)$ size of frame **q-1**, centred at the same spatial location as U_q and U_{q-1} is subblock from **SA**, where d_m is the maximum displacement allowed in either direction in integer number of pixel.

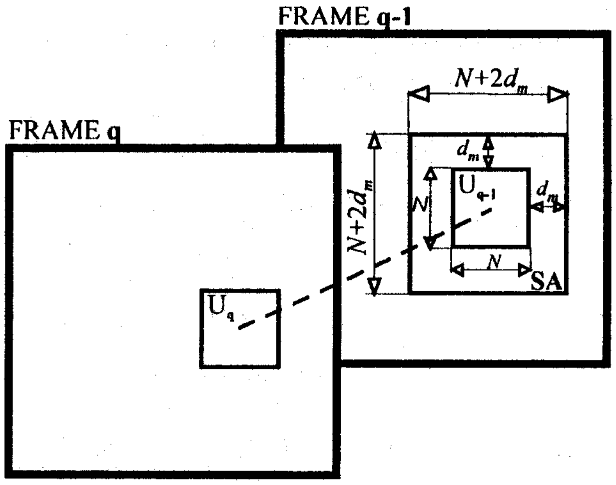


Fig.4 Positions of subblocks $U_{q(q-1)}$ and SA at the frames **q (q-1)**

4.1 Motion estimation algorithms with use of 1D-IRT

Let K_r and K_c are matrices of states computed by row and by column of subblock respectively.

STEP 1: Compute $K_{r(q)}$, $K_{c(q)}$ for subblock U_q.

STEP 2: Compute $K_{r(q-1)}$, $K_{c(q-1)}$ for subblock U_{q-1}.

STEP 3: Compute matching criterion

$$\sigma_{row}(u,v)=\sum_{r=0}^{\mu}\sum_{i=0}^{N-1}\sum_{j=0}^{N/2-1}\left(k_{r(k)}(i,j)\oplus k_{r(k-1)}(i,j)\right)$$

$$\sigma_{col}(u,v)=\sum_{r=0}^{\mu}\sum_{i=0}^{N-1}\sum_{j=0}^{N/2-1}\left(k_{c(k)}(i,j)\oplus k_{c(k-1)}(i,j)\right).$$

$$u,v\in\langle -d_m,d_m\rangle \quad (7)$$

Repeat steps 2, 3 for every possible positions (u,v) of subblock U_{q-1} in subblock SA ($(2d_m+1)^2$ cycles), where $\oplus$ denotes bit-by-bit modulo 2 addition.

STEP 4: The desired vector of motion correspond to the position (u_0, v_0) of subblock U_{q-1} with minimal value of $\sigma(u, v)$.

Modifications of the algorithm - I

μ -number of used columns of matrix K

$\mu\in\{0, 1, \dots, n-1\}$

Modifications of the algorithm - II

$$4a,\ (u_0,v_0)\in\{u,v\};\sigma_{row}(u_0,v_0)=\min(\sigma_{row}(u,v))$$

$$4b,\ (u_0,v_0)\in\{u,v\};\sigma_{col}(u_0,v_0)=\min(\sigma_{col}(u,v))$$

$$4c,\ (u_0,v_0)\in\{u,v\};\sigma_{row}(u_0,v_x)=\min(\sigma_{row}(u,v))\wedge \wedge\,\sigma_{col}(u_x,v_0)=\min(\sigma_{col}(u,v))$$

$$4d,\ (u_0,v_0)\in\{u,v\};\sigma_{r+c}(u_0,v_0)=\min(\sigma_{r+c}(u,v)), \quad (8)$$

where $\sigma_{r+c}(u,v)=\sigma_{row}(u,v)+\sigma_{col}(u,v)$ (9)

4.2 Motion estimation algorithm with use of 2D-IRT

STEP 1: Compute first set of $K_p^{(r)}$ (i.e. $K_{1(q)}^{(0)}, K_{2(q)}^{(0)}, K_{3(q)}^{(0)}, K_{4(q)}^{(0)}$) for block U_q.

STEP 2: Compute first set of $K_p^{(r)}$ (i.e. $K_{1(q-1)}^{(0)}, K_{2(q-1)}^{(0)}, K_{3(q-1)}^{(0)}, K_{4(q-1)}^{(0)}$) for block U_{q-1}.

STEP 3: Compute matching criterion

$$\sigma(u,v) = \sum_{p=1}^{\tau} \sum_{i=0}^{N/2-1} \sum_{j=0}^{N/2-1} \left(k_{p(q)}^{(0)}(i,j) \oplus k_{p(q-1)}^{(0)}(i,j) \right)$$

$$u,v \in \langle -d_m, d_m \rangle \qquad (10)$$

Repeat steps 2, 3 for every possible positions (u,v) of subblock U_{q-1} in subblock SA ($(2d_m+1)^2$ cycles), where $\oplus$ denotes bit-by-bit modulo 2 addition.

Modifications of the algorithm

3a, $\tau = 1$ (K1)

3b, $\tau = 2$ (K1, K2)

3c, $\tau = 3$ (K1, K2, K3)

3d, $\tau = 4$ (K1, K2, K3, K4)

STEP 4: The desired vector of motion corresponds to the position (u_0, v_0) of subblock U_{q-1} with minimal value of $\sigma(u,v)$, i.e.

$$(u_0, v_0) \in \{u,v\}; \sigma(u_0,v_0) = \min(\sigma(u,v)) \qquad (11)$$

5. Experiments and results

The methods mentioned above was simulated on personal computer. The results was found using a frames of 256 x 256 pixels quantized uniformly to 8 bits. The simulation was carried out with two types pictures. First picture "Talker" represented a talking person (picture type portrait with less details) and second picture "Printed text on textured background" represented a non-uniformly illuminated textured background, on which was translated a white area with printed text (picture with a lot of details). In both occasions the maximal presupposed displacement was $d_m = 4$.

The effect of motion estimation can be indicated by the improvement in the signal-to-noise ratio (**SNR**). We define

$$\mathrm{SNR} = 10\log_{10} \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} \left[\bar{x} - x(i,j)\right]^2}{\sum_{i=1}^{M} \sum_{j=1}^{N} \left[x(i,j) - x''(i,j)\right]^2}, \qquad (12)$$

where $\bar{x}$ is mean value of frame **q**, $x(i,j)$ are values of pixels of the frame and $x''(i,j)$ are values of reconstructed frame or frame **q-1**. The frame reconstruction means subblock settling of frame **q-1** by found movement vectors. The experimental results for the subblock size 4 x 4, 8 x 8, and 16 x 16 pixels, maximum displacement d_m=4 are shown in Fig.5-9, where pointed area called **Interframe** represents the level of SNR calculated from (12) for frames **q** and **q-1**, curve 2D-log represents results obtained from 2D-log method, curves designed with symbols **4a**, **4b**, **4c**, **4d** represent results obtained from 1D-IRT method for modification **4a-4d**, and curves designed with symbols **K1**, **K2**, **K3**, **K4** represent results obtained from 2D-IRT method using first set of matrices ($K_1^{(0)}, K_2^{(0)} K_3^{(0)}, K_4^{(0)}$) from system $K_p^{(r)}$.

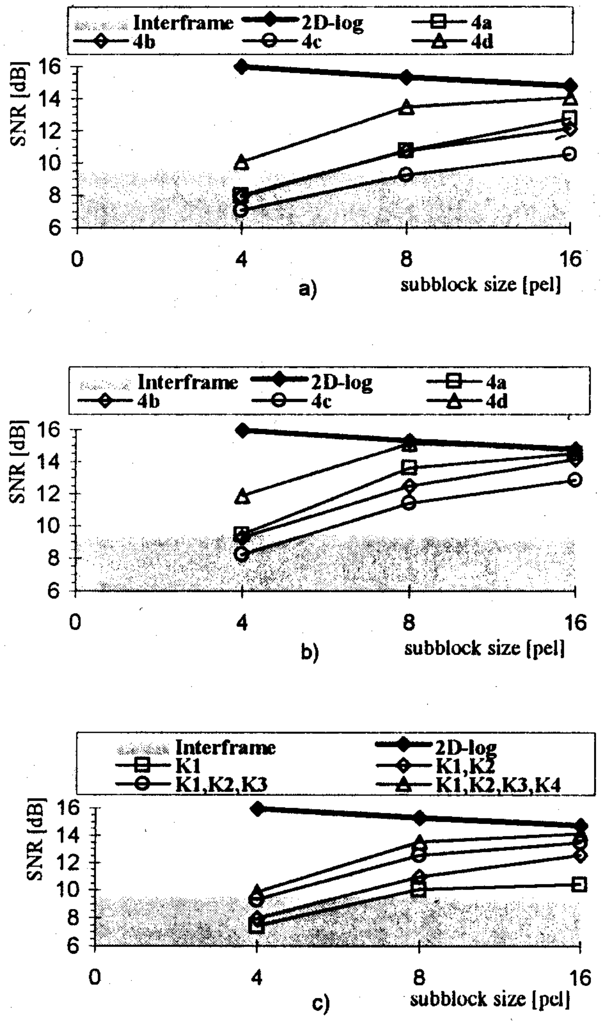


Fig.5: Experimental results for sequence "**Talker**"; **a)** the chart for 1D-IRT method with μ=1, **b)** the chart for 1D-IRT method with μ=2, **c)** the chart for 2D-IRT method.

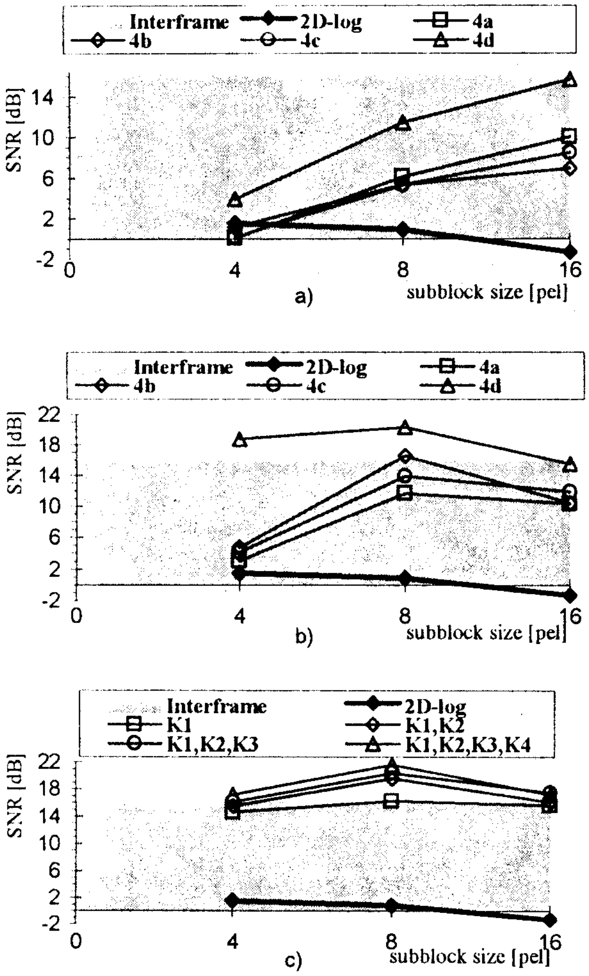


Fig.6: Experimental results for sequence "**Printed text on textured background**"; **a)** the chart for 1D-IRT method with μ=1, **b)** the chart for 1D-IRT method with μ=2, **c)** the chart for 2D-IRT method.

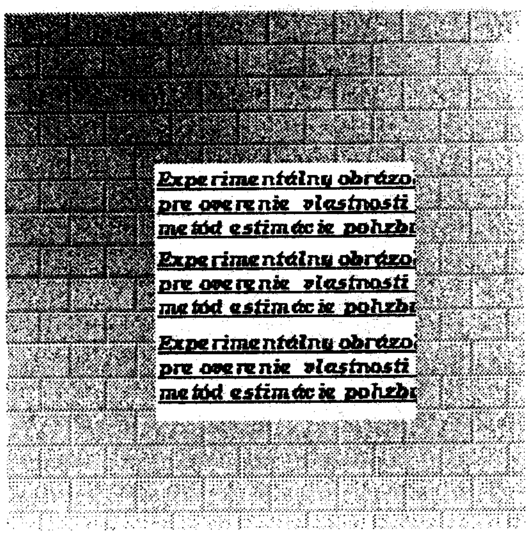
a)

b)

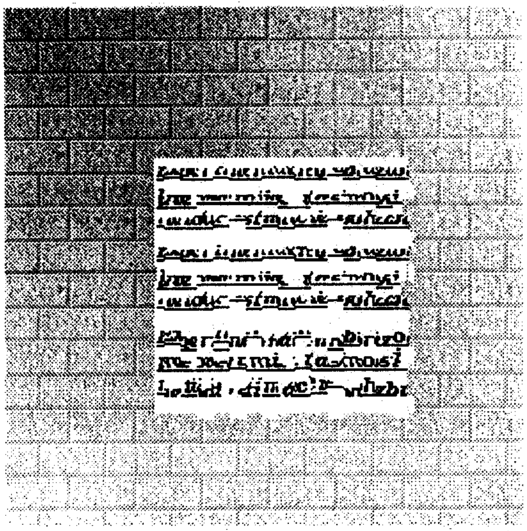
c)

d)

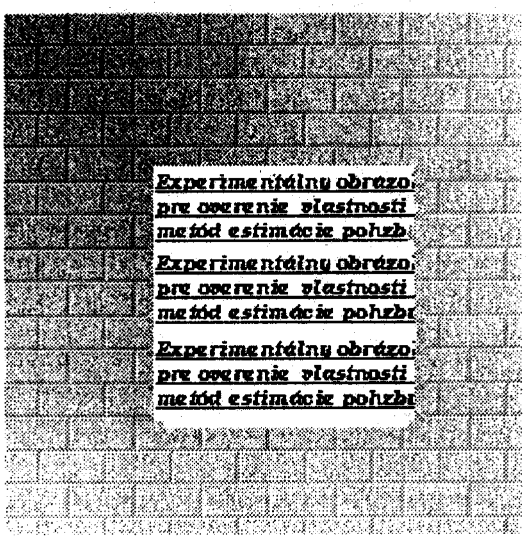
Fig.7: Experimental results for sequence "**Talker**"; **a)** the frame **q** **b)** reconstructed frame (2D-log method, subblock size =8, d_m=4); **c)** reconstructed frame (1D-IRT method, modification 4d, subblock size =8, d_m=4 and μ=2); **d)** reconstructed frame (2D-IRT method, modification K1-K4, subblock size=8, d_m=4).



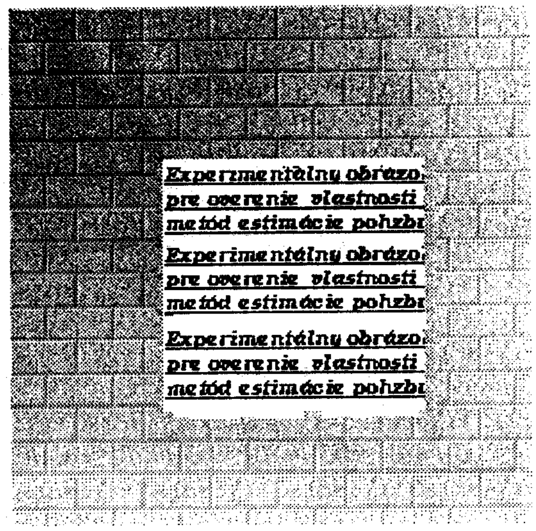
a)



b)

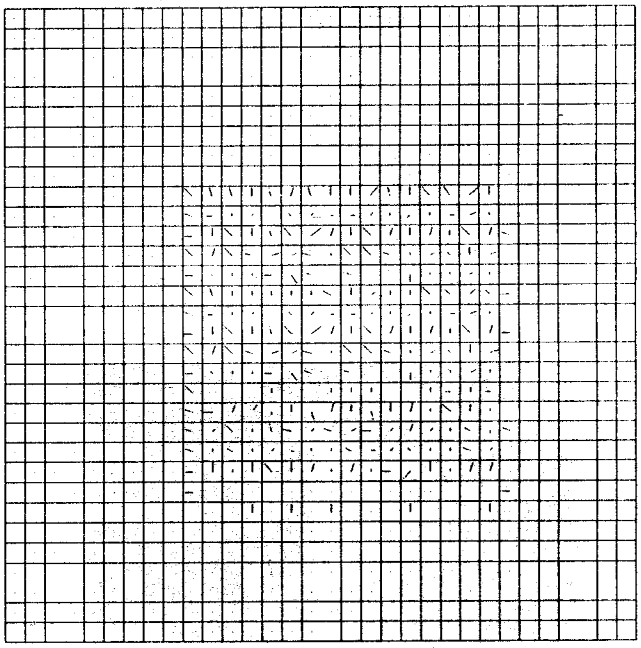


c)

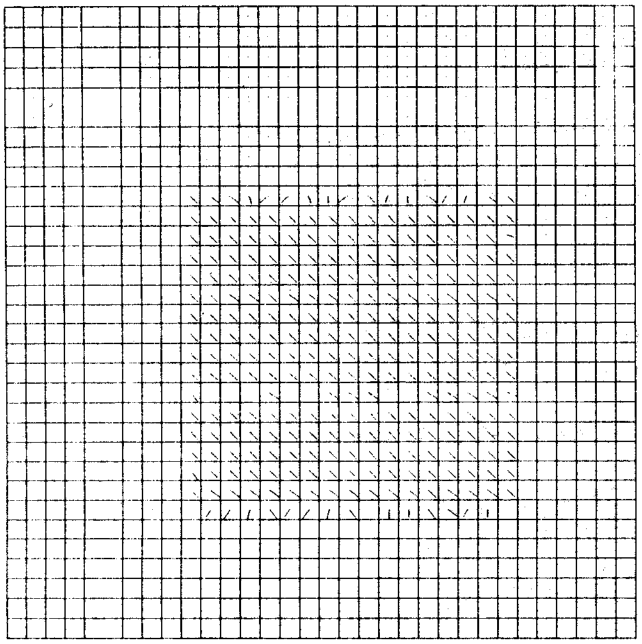


d)

Fig.8: Experimental results for sequence "**Printed text on textured background**"; **a)** the frame **q**; **b)** reconstructed frame (2D-log method, subblock size =8, d_m=4); **c)** reconstructed frame (1D-IRT method, modification 4d, subblock size =8, d_m=4 and μ=2); **d)** reconstructed frame (2D-IRT method, modification K1-K4, subblock size=8, d_m=4).



a)



b)

Fig.9: Motion vectors for sequence "**Printed text on textured background**"; a) 2D-log method (subblock size = 8, d_m= 4), b) 1D-IRT method (modification 4d, subblock size = 8, d_m= 4 and μ=2).

6. Conclusion

The experimental results indicate that proposed methods give for picture with less details slightly worse or comparable results as 2D-log method. Significantly better results are obtained for the proposed methods for pictures with a lot of details ("Printed text on textured background") and it is very interesting that its efficiency in contrast to 2D-log method is increasing with increasing size of subblock. There is also evidence for increasing of efficiency with the number of used matrices $\boldsymbol{K}_p^{(r)}$ for 2D-IRT method and increasing of efficiency for **4d** modification of 1D-IRT method.

References

[1] REITBOECK, H. - BRODY, T. P.: A Transformation with Invariance Under Cyclic Permutation for Applications in Pattern Recognition. Inf. and Control, Vol.15, 1969, 130-154.

[2] BURKHARDT, H.: Transformationen zur lageinvarianten Merkmalgewinnung. Fortschrittsbericht. (Reihe 10, Nr.7), VDI-Verlages, Düsseldorf, 1979.

[3] TURÁN, J.- KŐVESI, L.- KŐVESI, M.: CAD System for Pattern Recognition and DSP with Use of Fast Translation Invariant Transforms. J. on Communications, Vol. XLV, 1994, 85-89.

[4] TURÁN, J.- CHMÚRNY, J.: Two - Dimensional Fast Translation Invariant Transforms and Their Use in Robotics. Electronic Horizon, Vol. 15, No. 5, 1984, 211-220.

[5] TURÁN, J. - CHMÚRNY, J.: Two - Dimensional Inverse R-Transform. Computers and Artificial Intelligence, Vol.2, No.5, October 1983, 473-477.

[6] FANG, M.: Class of Invertible Shift Invariant Transform. Signal Processing. Vol.23, No.4, April 1991, 35-44.

[7] FANG, M. - HÄUSLER, G.: An Invertible Rapid Transform. Proc. 12-th DAGM-Symposium, 1990, 201-208.

[8] GAMEC,J.-TURÁN,J.: Motion Estimation with Use of 1D-Inverse Rapid Transform. In: Intelligent Terminals and Source and Channel Coding, COST 229, Budapest, Sept. 7.-9., 1993, 213-218.

[9] TURÁN, J. - GAMEC, J.: Application of 1D-Inverse Rapid Transform in Picture Coding. In: Intelligent Terminals and Source and Channel Coding, COST 229, Budapest, Sept. 7.-9., 1993, Hungary, 61-68.

[10] GAMEC, J. - TURÁN, J.: Inverse Rapid Transform and Motion Analysis. Proc. of Workshop COST 229, Bayona-Vigo, Spain, Oct. 1994.

[11] FANG, M.: A New Method for Fast Shift Determination. Proc. 12-th DAGM-Symposium, 1990, 240-247.

[12] WAGH. M. D. - KANETKAR, S. V.: A Class of Translation Invariant Transforms. IEEE Trans. on Acoustic, Speech and Signal Proc., Vol. ASSP-25, No. 3, 1977, 203-205.

[13] WANG, P. O. - SCHIAU, R. C.: Machine Recognition of Printed Chinese Characters via Transformation Algorithms. Pattern Recognition, Vol. 5, 1973, 303-321.

[14] NASARIMHAN, M. A. - DEVARAJAN, V. - RAO, K. R.: Simulation of Alphanumeric Machine Print Recognition. IEEE Trans. Vol. SMC-10, 1980, 270-275.

[15] SCHÜTTE, S. - FRYDRYCHOWICZ, J. - SCHRŐDER, J.: Scene Matching with Translation Invariant Transforms. 5-ICPR, Miami, USA, 1980, 195-198.

[16] FANG, M. - HÄUSLER, G.: Modified Rapid Transform. Applied Optics, Vol. 28, No. 6, 1989, 1257-1262.

[17] TURÁN, J.: Recognition of Printed Berber Characters Using Modified Rapid Transform. J. on Communication, Vol. XLV, 1994, 24-27.

[18] TURÁN, J. - ALTHŐFER, K.: A Novel System for a 3D Acoustic Object Recognition Based on the Modified Rapid Transform. J. El. Eng., Vol. 46, No. 8, 1995, 265-269.

About authors...

Ján GAMEC (Ing., CSc.), was born in Stul'any, Slovakia in 1960. He received an Ing (MSc) degree in Radiotechnics with honours from the Technical University, Košice, Slovakia in 1985. He received a CSc (PhD) degree in radioelectronics from the Technical University, Košice, Slovakia, in 1995. Since August 1985, he has been at the Technical University of Košice, as a PhD student and assistant professor of electronics and information technology. His research interests include digital image processing.

Ján TURÁN (Prof., Ing., RNDr., DrSc.), was born in Šahy, Slovakia in 1951. He received an Ing. (MSc.) degree in physical engineering with honours from the Czech