Chapter 4
Invariant Object Representation Based on Inverse Pyramidal Decomposition and Modified Mellin-Fourier Transform

R. Kountchev, S. Rubin, M. Milanova, and R. Kountcheva

Abstract In this work is presented one new method for invariant object representation based on the Inverse Pyramidal Decomposition (IPD) and modified Mellin-Fourier Transform (MFT). The so prepared object representation is invariant against 2D rotation, scaling, and translation (RST). The representation is additionally made invariant to significant contrast and illumination changes. The method is aimed at content-based object retrieval in large databases. The experimental results obtained using the software implementation of the method proved its efficiency. The method is suitable for various applications, such as detection of children sexual abuse in multimedia files, search of handwritten and printed documents, 3D objects, represented by multi-view 2D images, etc.

R. Kountchev (✉)
Department of Radio Communications, Technical University of Sofia, Bul. Kl. Ohridsky 8,
Sofia, 1000, Bulgaria
e-mail: rkountch@tu-sofia.bg

S. Rubin
Space and Naval Warfare Systems Center San Diego (SSC-PAC), San Diego, CA, USA
e-mail: stuart.rubin@navy.mil

M. Milanova
Computer Science Department, UALR, 2801 South University Avenue, Little Rock,
AR 72204, USA
e-mail: mgmilanova@ualr.edu

R. Kountcheva
T&K Engineering, Mladost 3, Pob.12, Sofia, 1712, Bulgaria
e-mail: kountcheva_r@yahoo.com

T. Özyer et al. (eds.), *Information Reuse and Integration in Academia and Industry*,
DOI 10.1007/978-3-7091-1538-1_4, © Springer-Verlag Wien 2013

Introduction

One of the most important tasks, which the creators of contemporary computer visual systems have to solve, is related to objects description, which ensures their exact and reliable classification. Objects description should satisfy contradictory requirements [1, 2]: to be invariable to object rotation, translation and scaling; lighting (respectively – contrast) changes; noises in the image; low intra-class and high inter-class dispersion; maximum compactness of the description and low computational complexity of the operations performed.

The objects descriptions could be divided into two basic groups: local and global. The first group represents the local features of the objects, related to the structure of the visible part of their surface (salient points, contours, edges, texture, etc.) [3, 4]. These descriptions are not resistant enough against noises and have high dimensionality and computational complexity. The second group comprises characteristics, which represent the general qualities of the visible surface, related for example, to object's geometrical parameters (invariable moments), coefficients of the contour lines' Fourier decomposition, global brightness histograms, coefficients of the generalized Hough transform, coefficients of the Mellin-Fourier transform, etc. [5–8]. The advantages of these descriptions are their insensibility to noises and the big compactness, but their computational complexity is relatively high.

The basic methods for invariant object representation with respect to 2D rigid transforms (combinations of rotation, scaling, and translation, RST) are given in significant number of scientific publications. Accordingly, 2D objects in the still grayscale image are depicted by descriptors of two basic kinds: "shape boundary" and "region". To the first kind (shape boundary) are assigned the chain codes, Fourier descriptors; Generalized Hough Transform and Active Shape Model [7]. The skeleton of a shape can be derived by the Medial Axis Transform. To the second kind (region) are assigned some geometric characteristics, such as for example: area, perimeter, compactness, dispersion, eccentricity, etc., zero- and first-order statistical moments, centre of gravity, normalized central moments, seven rotation-invariant moments, Zernike polynomial rotation- and scale-invariant, affine transform invariant in respect to position, rotation and different scales along the coordinate axes, co-occurrence texture descriptor, etc.

The histogram descriptor is proved to be robust to changes of object's rotation, scale, and partial changes in the viewing direction. The structural information however is lost in the histogram. To solve this problem, the combination of Discrete Wavelet Transform (DWT) or Discrete Fourier Transform (DFT) with the feature extraction method is proposed. For the extraction of the rotation-scale-translation (RST) – invariant features are developed descriptors, based on the log-polar transform (LPT) used to convert rotation and scaling into translation [8] and on the 2D Mellin-Fourier Transform (2D-MFT) [9]. As it is known, the modules of the spectrum coefficients, obtained using the 2D-MFT, are invariant with respect to the RST-transforms of the 2D objects in the image. The basic problem for the creation of the RST-invariant descriptors, in this case is the large number of spectrum coefficients, which have to be calculated [10, 11]. With regard to the necessity to reduce their number, and respectively – the time needed for the calculation without decreasing the objects description accuracy, should be solved significant number of problems, regarding the choice of the most informative MFT coefficients and the way of creating the corresponding vector descriptor.

In this work is offered new algorithm for global description of pre-segmented objects in a halftone image. This algorithm differs from the known algorithms in the same group: it requires lower number of computations when compared to the famous MFT and permits multi-layer access which enhances the search by content in indexed image databases. The new algorithm is based on the methods of the inverse pyramid decomposition [12] and of the truncated modified discrete MFT [13, 14], which are combined.

This work is arranged as follows: in section "Invariant Object Representation with Pyramid of Coefficients Based on the Truncated Modified Mellin-Fourier Transform" is described the method for invariant object representation based on pyramid of coefficients, calculated using the truncated Mellin-Furier transform; in section "Search by Content of Closest Objects in Image Databases" is given the new approach for search-by-content of closest object in image database; in section "Fast Search of Closest Vector in the Image Database" – the fast search of closest vector in the image database, in section "Experimental Results" are shown some experimental results, and in section "Conclusions" are given the Conclusions.

Invariant Object Representation with Pyramid of Coefficients Based on the Truncated Modified Mellin-Fourier Transform

The algorithm for 2D object representation is aimed at the preparation of the vector description of the segmented object, framed by a square window.

The description should be invariant to 2D rotation (R), scaling (S), translation (T) and contrast (C) changes. As a basis for the RSTC description is used the discrete 2D Modified MFT (2D-MMFT). As it is known, the MFT comprises DFT, Log-pol transform (LPT) and DFT again. In order to provide multi-layer search-by-content in the image database, the 2D-MMFT coefficients are arranged in a pyramid, called Inverse Pyramid Decomposition (IPD).

The algorithm, presented below, is aimed at digital halftone images, and comprises the following stages:

For the initial (lowest) IPD level:

Step 1. The pixels B(k,l) of the original halftone image of size M × N are transformed into bi-polar:

$$L(k,l) = B(k,l) - (B_{\max} + 1)/2 \tag{4.1}$$

for $k = 0, 1, \ldots, M-1$ and $l = 0, 1, .., N-1$, where $B_{\max} = 255$ is the maximum value in the pixel quantization scale.

Step 2. The image is processed with 2D Discrete Fourier Transform (2D-DFT). The Fourier matrix is of size $n \times n (n = 2^m)$. The value of n defines the size of the window, used to select the object image. For the invariant object representation are used the complex 2D-DFT coefficients, calculated in accordance with the relation:

$$\mathrm{F(a, b)} = \sum_{\mathrm{k=0}}^{n-1} \sum_{\mathrm{l=0}}^{n-1} \mathrm{L(k, l)} \exp\{-\mathrm{j}[(2\pi/\mathrm{n})(\mathrm{ka} + lb)]\} \tag{4.2}$$

for $a = 0, 1, .., n - 1$ and $b = 0, 1, .., n - 1$.

The transform comprises two consecutive operations: one-dimensional transform of the pixels L(k,l), first – for the rows and after that – for the columns of the object image. Since

$$\exp\{-j\,[2\pi(\mathrm{lb/n})]\} = \cos[2\pi(\mathrm{lb/n})] - \mathrm{j}\sin[2\pi(\mathrm{lb/n})]$$

and

$$\exp\{-j\,[2\pi(\mathrm{ka/n})]\} = \cos[2\pi(\mathrm{ka/n})] - \mathrm{j}\sin[2\pi(\mathrm{ka/n})],$$

the 2D-DFT is performed as two consecutive one-dimensional DFTs:

- For the fixed values of $k = 0, 1, .., n-1$ and using the 1D-Fast Fourier Transform (1D-FFT) are calculated the intermediate spectrum coefficients:

$$\begin{aligned} F(k, b) &= \sum_{l=0}^{n-1} L(k, l) \exp\{-j\,[2\pi(\mathrm{lb/n})]\} = \\ &= \sum_{\mathrm{l=0}}^{\mathrm{n-1}} \mathrm{L(k, l)} \cos[2\pi(\mathrm{lb/n})] - j \sum_{\mathrm{l=0}}^{\mathrm{N-1}} \mathrm{L(k, l)} \sin[2\pi(\mathrm{lb/n})] \end{aligned} \tag{4.3}$$

- For $b = 0, 1, \ldots, n - 1$ and using the 1D-FFT again, are calculated the final Fourier coefficients:

$$\begin{aligned} F(a, b) &= \sum_{k=0}^{n-1} F(k, b) \exp\{-j\,[2\pi(\mathrm{ka/n})]\} = \sum_{\mathrm{k=0}}^{\mathrm{n-1}} F(k, b) \cos[2\pi(\mathrm{ka/n})] - \\ &-j \sum_{\mathrm{k=0}}^{\mathrm{n-1}} F(k, b) \sin[2\pi(\mathrm{ka/n})] = \mathrm{A_F(a, b)} - \mathrm{jB_F(a, b)} \end{aligned} \tag{4.4}$$

where A_F (a, b) and B_F (a, b) are the real and the imaginary components of F(a, b) correspondingly.

Step 3. The Fourier coefficients are then centered in accordance with the relation:

$$F_0(a, b) = F\left(a - \tfrac{n}{2}, b - \tfrac{n}{2}\right) \text{ for a, b} = 0, 1, .., \mathrm{n} - 1. \tag{4.5}$$

Step 4. For the next operations some of the Fourier coefficients are retained in accordance with the rule:

$$\mathrm{F_{0R}(a,b)} = \begin{cases} \mathrm{F_0}(a,b), if(a,b) \in \textit{retained region}; \\ 0 \qquad - \qquad \textit{in all other cases}. \end{cases} \tag{4.6}$$

The retained coefficients' area is a square with a side H $\leq$ n, which envelopes the centre (0,0) of the spectrum plane (H – even number). For H $<$ n, and $a, b = -(H/2), -(H/2) + 1, \ldots, -1, 0, 1, \ldots, (H/2) - 1$, this square contains low-frequency coefficients only.

Step 5. The modules and phases of coefficients $\mathrm{F_{0R}(a,b)} = \mathrm{D_{F_{0R}}}(\mathrm{a}, b)\mathrm{e}^{\mathrm{j}\varphi_{\mathrm{F0R}}(\mathrm{a},b)}$ are calculated:

$$\mathrm{D_{F_{0R}}(a,b)} = \sqrt{[A_{F_{0R}}(a,b)]^2 + [B_{F_{0R}}(a,b)]^2} \tag{4.7}$$

$$\varphi_{\mathrm{F_{0R}}}(\mathrm{a,b}) = arctg[B_{F_{0R}}(a,b)/A_{F_{0R}}(a,b)] \tag{4.8}$$

Step 6. The modules $\mathrm{D_{F_{0R}}(a,b)}$ of the Fourier coefficients $\mathrm{F_{0R}(a,b)}$ are normalized in accordance with the relation:

$$\mathrm{D(a,b)} = p \,\ln D_{F_{0R}}(a,b) \tag{4.9}$$

where p = 64 is the normalization coefficient.

Step 7. The coefficients D(a, b) are processed with Log-Polar Transform (LPT). The centre (0,0) of the polar coordinate system (ρ, θ) coincides with the centre of the image of the Fourier coefficients' modules D(a, b) (in the rectangular coordinate system). The transformation of coefficients D(a, b) from the rectangular (a, b) into the polar (ρ, θ) coordinate system is performed changing the variables in accordance with the relations:

$$\rho = \log\sqrt{\mathrm{a}^2 + \mathrm{b}^2}, \quad \theta = \mathrm{arctg(b/a)} \tag{4.10}$$

The coordinate change from rectangular into polar is quite clear in the continuous domain, but in the discrete domain the values of ρ and θ should be discrete as well.

Since a and b can only have discrete values in the range $a, b = -(H/2), \ldots, -1, 0, 1, \ldots, (H/2) - 1$, some of the coefficients $\mathrm{D}(\rho, \theta)$ will be missing. At the end of the transform, the missing coefficients $\mathrm{D}(\rho_i, \theta_i)$ are interpolated using the closest neighbors D(a, b) in the rectangular coordinate system (a, b) in horizontal or vertical direction (zero-order interpolation).

The number of discrete circles in the polar system with radius ρ_i is equal to the number of the discrete angles θ_i for i =, 2, .., H. The size (in rectangular coordinates) of the side of the square H inscribed in the LPT matrix is calculated so, that to ensure maximum part of the coefficients to be transferred without change.

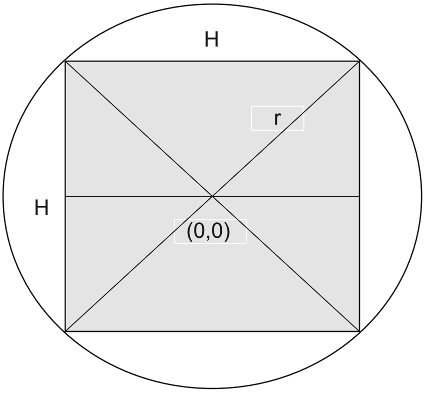


Fig. 4.1 Geometric relations between *r* and *H*

For this, the LP transform is modified in accordance with Fig. 4.1, calculating the radius of the circumscribed circle in correspondence to the relation:

$$\mathrm{r} = (\sqrt{2/2})\mathrm{H} \tag{4.11}$$

The smallest step $\Delta\rho$ between two concentric circles (the most inside) is calculated as follows:

$$\Delta\rho = r^{(1/H)}. \tag{4.12}$$

The square, with side H, which defines the retained coefficients' area.

As a result, for the discrete radius ρ_i and angle θ_i for each circle are obtained the relations:

$$\rho_i = (\Delta\rho)^i = r^{i/H} \text{ for i} = 1, 2, .., \mathrm{H}, \tag{4.13}$$

$$\theta_i = (2\pi/H)i \text{ for } i = (-H/2), ., 0, ., (-H/2) - 1. \tag{4.14}$$

Thus, instead of the logarithmic relation used in the famous LP transform to set the values of the magnitude bins (radiuses), in Step 7 here is used the operation "rising on a power". The so modified LP transform we called *Exponential Polar Transform* (EPT).

After the EPT and the interpolation of the D(a, b) coefficients is obtained one new, second matrix, which contains the coefficients D(x,y) for x, y $= 0, 1, 2, .., \mathrm{H} - 1$.

Step 8. The second 2D-DFT is performed for the matrix with coefficients D(x,y), in accordance with the relation:

$$\mathrm{S(a, b)} = \frac{1}{H^2}\sum_{x=0}^{H-1}\sum_{y=0}^{H-1} \mathrm{D_I}(x, y)\exp\{-\mathrm{j}[(2\pi/H)(x\mathrm{a} + yb)]\} \tag{4.15}$$

for $a = 0, .., H - 1$ and $b = 0, ..H - 1$.

The second 2D-DFT is performed in correspondence with Eqs. 4.3 and 4.4, applying consecutively the 1D-DFT on the rows of the matrix [D] first, and then – on the columns of the intermediate matrix obtained.

Step 9. The modules of the complex coefficients S(a, b) are then calculated:

$$\mathrm{D_S(a,b)} = \sqrt{[A_S(a,b)]^2 + [B_S(a,b)]^2} \tag{4.16}$$

where A_S(a, b) and B_S(a, b) are the real and the imaginary component of S(a, b) correspondingly.

With this operation the Modified MFT is finished.

The processing then continues in the next step with one more operation, aimed at achieving the invariance against contrast changes. In result is obtained the RSTC invariant object representation.

Step 10. The modules $\mathrm{D_S}$(a, b) of the Fourier coefficients S(a, b) are normalized:

$$\mathrm{D_{S_0}(a,b)} = B_{\max}(D_S(a,b)/D_{S\max}(a,b)), \tag{4.17}$$

where $D_{S\max}(a,b)$ is the maximum coefficient in the matrix $[D_S(a,b)]$.

Step 11. The vector for the RSTC-invariant object representation is based on the use of coefficients $\mathrm{D_{S_0}}$(a, b) of highest energy in the amplitude spectrum 2D-MFT of size H × H. One example test image (Caltech database "Faces" [15]) and its corresponding 2D-MFT spectrum obtained with the software implementation of the method are shown correspondingly on Fig. 4.2a and b. For the extraction of the retained coefficients is used the mask, shown on Fig. 4.2c (the part, colored in yellow, corresponds to the area of the complex-conjugated coefficients,). The shape of the mask approximates the area, where the energy of the mean 2D-MFT spectrum is concentrated.

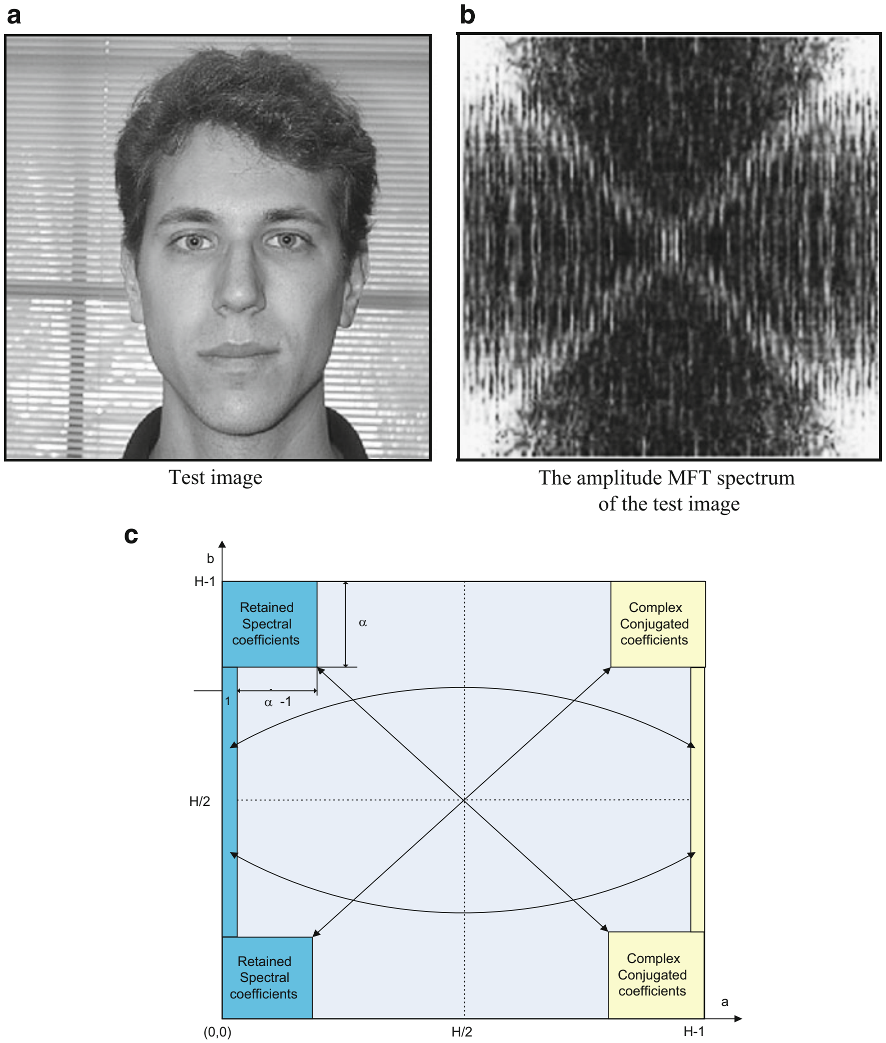
The parameter α of the mask defines the number of retained coefficients in correspondence with the relation:

$$R = 2\alpha^2 + H - 2\alpha = 2\alpha(\alpha - 1) + H \tag{4.18}$$

The components v_m for m = 1, 2, .., R of the corresponding RSTC-invariant vector: $\mathbf{V}_0 = [v_{01}, v_{02}, .., v_{0R}]^T$ are defined by coefficients $\mathrm{D_{S_{0R}}}$(a, b), arranged as one-dimensional massif after lexicographic tracking of the 2D-MFT spectrum in the mask area, colored in blue.

The vector $\mathbf{V}_0$ is the RST-invariant description of the processed image for the initial (zero) IPD level. The block diagram of the algorithm for vector calculation in a square window of size n × n is shown on Fig. 4.3.

For the calculation of the vector $\mathbf{V}_1$ for the next IPD level (one) of the processed image is performed inverse Modified Mellin-Fourier transform for the coefficients $\mathrm{D_S}$(a, b). In result is obtained the approximation $\hat{L}(k,l)$ of the processed image. For this is performed the following:



Test image

The amplitude MFT spectrum
of the test image

Mask of retained coefficients in the 2D-MFT spectrum

Fig. 4.2 The mask of the retained coefficients in the amplitude 2D-MFT spectrum of the test image

Step 1. The denormalized coefficients are calculated:

$$D'_S(a,b) = [D_{S_0}(a,b)D_S(a,b)_{\max}]/B_{\max}. \tag{4.19}$$

Step 2. The complex coefficients are then calculated:

$$\mathrm{S}'(\mathrm{a},\mathrm{b}) = \mathrm{D}'_\mathrm{S}(\mathrm{a},\mathrm{b})\mathrm{e}^{-\mathrm{j}\varphi_\mathrm{S}(\mathrm{a},\mathrm{b})} \tag{4.20}$$

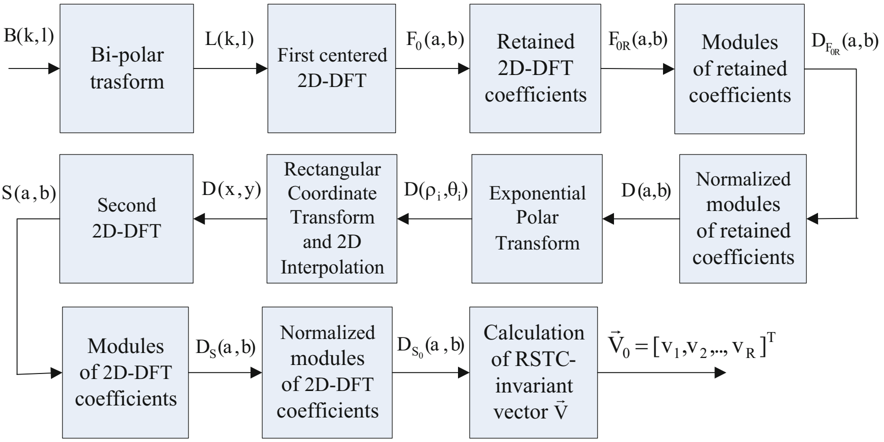

B(k,l)
Bi-polar trasform
L(k,l)
First centered 2D-DFT
$F_0(a,b)$
Retained 2D-DFT coefficients
$F_{0R}(a,b)$
Modules of retained coefficients
$D_{F_{0R}}(a,b)$
Normalized modules of retained coefficients
D(a,b)
Exponential Polar Transform
$D(\rho_i,\theta_i)$
Rectangular Coordinate Transform and 2D Interpolation
D(x,y)
Second 2D-DFT
S(a,b)
Modules of 2D-DFT coefficients
$D_S(a,b)$
Normalized modules of 2D-DFT coefficients
$D_{S_0}(a,b)$
Calculation of RSTC-invariant vector $\vec{V}$
$\vec{V}_0=[v_1,v_2,..,v_R]^T$

Fig. 4.3 Block diagram of the algorithm for calculation of the R-dimensional RSTC-invariant vector for object description in a *square* window of size n × n, through direct MMFT

Step 3. The so calculated coefficients are processed with inverse 2D-DFT:

$$\mathrm{D}'(\mathrm{x},\mathrm{y})=\sum_{\mathrm{a}=0}^{\mathrm{H}-1}\sum_{\mathrm{b}=0}^{H-1}S'(a,b).\exp\left\{\mathrm{j}\left[\frac{2\pi}{\mathrm{H}}(\mathrm{xa}+\mathrm{yb})\right]\right\} \tag{4.21}$$

for x, $\mathrm{y}=0,1,2,..,\mathrm{H}-1$.

Step 4. Then coefficients $\mathrm{D}'(\mathrm{x},\mathrm{y})$ are processed with inverse EPT, after replacing the variables (x,y) by (a,b) correspondingly:

$$\begin{aligned}\mathrm{a_i}(\rho_i,\theta_i)&=[\mathrm{x}(\rho_i,\theta_i)]\cos[\mathrm{y}(\rho_i,\theta_i)],\\ \mathrm{b_i}(\rho_i,\theta_i)&=[\mathrm{x}(\rho_i,\theta_i)]\sin[\mathrm{y}(\rho_i,\theta_i)].\end{aligned} \tag{4.22}$$

In result are obtained the values of coefficients $\mathrm{D}'(\mathrm{a_i},\mathrm{b_i})$, from which after interpolation are restored the missing coefficients in the rectangular coordinate system (a, b). For this is used zero interpolation (each of the missing coefficients $\mathrm{D}'(\mathrm{a},\mathrm{b})$ is replaced by the closest one from the group of the existing coefficients in horizontal or vertical direction in the system $(\mathrm{a}_i,\mathrm{b}_i)$).

Step 5. The values of the interpolated coefficients $\mathrm{D}'(\mathrm{a},\mathrm{b})$ are denormalized in correspondence with the relation:

$$\hat{D}_{F_{0R}}(a,b)=\exp[\mathrm{D}'(\mathrm{a},\mathrm{b})/\mathrm{p}]. \tag{4.23}$$

Step 6. The retained complex coefficients are calculated:

$$\hat{\mathrm{F}}_{0\mathrm{R}}(\mathrm{a},\mathrm{b})=\hat{\mathrm{D}}_{\mathrm{F}_{0\mathrm{R}}}(\mathrm{a},\mathrm{b})\mathrm{e}^{-\mathrm{j}\varphi_{\mathrm{F}_{0\mathrm{R}}}(\mathrm{a},b)} \tag{4.24}$$

Step 7. The coefficients $\hat{\mathrm{F}}_{0\mathrm{R}}(\mathrm{a,b})$ are supplemented with zeros in accordance with the rule below:

$$\hat{\mathrm{F}}_0(\mathrm{a,b}) = \begin{cases} \hat{\mathrm{F}}_{0\mathrm{R}}(a,b), & if(a,b) \in \textit{retained region};\\ 0 & - \quad \textit{in all other cases}. \end{cases} \tag{4.25}$$

Step 8. Second, inverse 2D-DFT on the coefficients $\hat{\mathrm{F}}_0(a,\mathrm{b})$ is performed:

$$\hat{\mathrm{L}}(\mathrm{k,l}) = \sum_{a=-\frac{\mathrm{n}}{2}}^{\frac{\mathrm{n}}{2}-1} \sum_{b=-\frac{\mathrm{n}}{2}}^{\frac{\mathrm{n}}{2}-1} \hat{\mathrm{F}}_0(a,b)\exp\{j[2\pi/n(\mathrm{k}a+\mathrm{l}b)]\} \tag{4.26}$$

for $k = 0, 1, .., n-1$ and $l = 0, 1, .., n-1$.

As a result is obtained the approximated image with pixels $\hat{\mathrm{L}}(\mathrm{k,l})$ in the frame of the window of size n × n, which contains the object.

For the First IPD level is performed the following:

- The difference image $\mathrm{E}_0(\mathrm{k,l})$ is calculated:

$$E_0(k,l) = L(k,l) - \hat{L}(k,l) \text{ for k,l} \in \left(-\frac{n}{2}, \ldots 0 \ldots, \frac{n}{2}-1\right). \tag{4.27}$$

- The so obtained difference image is divided into four equal sub-images and each is after that processed with the already described direct Mellin-Fourier transform, following steps 2–8. The only difference is that in this case each sub-image is a square of size n/2.
- In a way, similar with Step 11, are calculated the corresponding RSTC vectors for the next (second) IPD level. The length of the vectors is four times smaller. The so calculated vectors are used to compose the general vector for the processed pyramid level – correspondingly $\mathbf{V}_1$ – whose length is equal with that of the vector $\mathbf{V}_0$, calculated for the zero pyramid level.

With this, the building of the two-level pyramid decomposition is finished. As a result are obtained the vectors $\mathbf{V}_0$ and $\mathbf{V}_1$, which are after that used for the search of the closest object in the image database (DB). Each image in the DB is represented by vectors, calculated in similar way for each of the pyramid levels, following the algorithm, presented above. For each IPD level is calculated a corresponding vector, which carries the information about the fine details, which represent the form of the object in the image.

The algorithm for calculation of RSTC-invariant vectors, presented above, could be generalized for IPD of $\mathrm{m} = \lg_2 \mathrm{n}$ levels. The block diagram of IPD-MMFT for $\mathrm{p} = 0, 1, .., \mathrm{m}-1$ is shown on Fig. 4.4. In each consecutive decomposition level is obtained the corresponding vector $\mathbf{V}_p$, which presents information for more and more small details, which represent the form of the object in the image L,

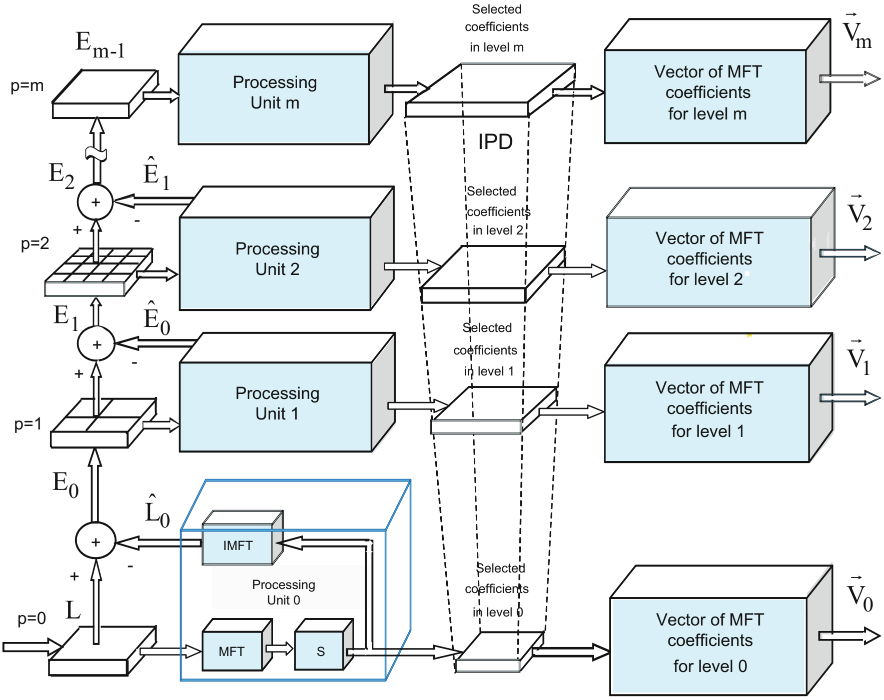


Fig. 4.4 Block diagram of the algorithm for calculation of the RST-invariant vectors $\mathbf{V}_0, \mathbf{V}_1, .., \mathbf{V}_m$, using IPD-MMFT of m levels

of size n × n. It is supposed, that the object is cropped from a larger image after corresponding preprocessing based on segmentation by color, form, etc.

Search by Content of Closest Objects in Image Databases

The search of closest objects in image databases (DB) for the image request is based on the detection of the minimum squared Euclidean distance (EUD) d between their RSTC vectors. For two R-dimensional vectors $\mathbf{V}_i$ and $\mathbf{V}_j$ this distance is defined by the relation:

$$d_E(\mathbf{V}_i, \mathbf{V}_j) = \sum_{m=0}^{R-1} [v_i(m) - v_j(m)]^2, \tag{4.28}$$

where $v_i(m), v_j(m)$ are the mth components of vectors $\mathbf{V}_i, \mathbf{V}_j$ for i $\neq$ j.

The decision for the image request classification, represented by the R-dimensional RSTC-vector $\mathbf{V}$ is taken on the basis of the image classes in the DB and their RSTC-invariant vectors:

$$\mathbf{V}_{\alpha_\beta}^{\beta} \in C_\beta \text{ for } \alpha_\beta = 1, 2, \ldots, \mathrm{P}_\beta \text{ and } \beta = 1, 2, \ldots., \mathrm{Q}; \tag{4.29}$$

P_β – the number of vectors in the DB;
Q – the number of image classes.

After that, a classification rule is applied, based on the K-nearest neighbors (k-NN) and "majority vote" algorithms [5]:

$$\mathbf{V} \in C_{\beta_0}, \ if \ \mathrm{d}_1\left(\mathbf{V}, \mathbf{V}_{\alpha_{\beta_1}}^{\beta_1}\right) \leq \mathrm{d}_2\left(\mathbf{V}, \mathbf{V}_{\alpha_{\beta_2}}^{\beta_2}\right) \leq \ldots. \leq \mathrm{d}_\mathrm{K}\left(\mathbf{V}, \mathbf{V}_{\alpha_{\beta_K}}^{\beta_K}\right) \tag{4.30}$$

where K is an odd number; $\mathrm{d}_\mathrm{k}\left(\mathbf{V}, \mathbf{V}_{\alpha_{\beta_k}}^{\beta_k}\right)$ is the squared Euclidian distance between vectors $\mathbf{V}$ and $\mathbf{V}_{\alpha_{\beta_k}}^{\beta_k}$ for k = 1, 2, ..., K; $\mathbf{V}_{\alpha_{\beta_k}}^{\beta_k}$ is the kth vector with index α_{β_k} from the class C_{β_k}, which is at minimum distance from the query $\mathbf{V}$ (the indices α_{β_k}, β_k and β_0 are in the ranges: $[1, P_{\beta_k}]$ – for α_{β_k}, and [1, Q] – for β_k and β_0). The class C_{β_0} of the vector $\mathbf{V}$ in Eq. 4.30 is defined by the most frequent value β_0 of the indices β_k of the vectors $\mathbf{V}_{\alpha_{\beta_k}}^{\beta_k}$:

$$\beta_0 = \max\{\mathrm{h}(\beta \mathrm{k})\} \text{ for } \beta_\mathrm{k} = 1, 2, .., \mathrm{Q} \text{ and } \mathrm{k} = 1, 2, .., \mathrm{K}. \tag{4.31}$$

Here $\mathrm{h}(\beta_k)$ is the histogram of the indices β_k, for the relations in Eq. 4.31.

For the enhancement of the object search, the smallest distance for the image request is calculated to a group of images, using the vectors defined for the zero level of the corresponding couple of pyramid decompositions. After that the smallest distance is calculated in the so selected group only, using the vectors for the next (first) pyramid level of the corresponding pyramids, etc.

Fast Search of Closest Vector in the Image Database

In order to perform the search of closest vector $\mathbf{V}_j$ (j = 1, 2, .., J) to the vector request $\mathbf{V}$ in the database (DB) with $J = Q \sum_{\beta=1}^{Q} P_\beta$ vectors could be used the approach in [16], based on the EUD between R-dimensional vectors:

$$d_E(\mathbf{V}, \mathbf{V}_j) = \sum_{m=0}^{R-1} [v(m) - v_j(m)]^2 = \|\mathbf{V}\|^2 + \|\mathbf{V}_\mathrm{j}\|^2 - 2\sum_{m=0}^{R-1} \mathrm{v(m)} \times \mathrm{v_j(m)}, \tag{4.32}$$

where $\|\mathbf{V}\|^2 = \sum_{m=0}^{R-1} [v(m)]^2$, $\|\mathbf{V}_\mathrm{j}\|^2 = \sum_{m=0}^{R-1} [v_j(m)]^2$.

The modules of vectors $\mathbf{V}$ and $\mathbf{V}_j$ in the equation above are not related to $d_E(\mathbf{V}, \mathbf{V}_j)$, but their scalar product is related as follows:

$$f(j) = \sum_{m=0}^{R-1} \mathrm{v(m)} \times \mathrm{v_j(m)}. \tag{4.33}$$

The vectors $\mathbf{V}_j$ (j = 1, 2, .., J) are calculated in advance for the images used as a training set for each image class and are stored as additional information (metadata) in the DB. In case, that the function f(j) has a maximum for some value of the variable j = j$_0$, the corresponding distance $d_E(\mathbf{V}, \mathbf{V}_{j0})$ is minimum. For the exact detection of j$_0$ is accepted the value, for which j$_0$ = min. The closest vector $\mathbf{V}_{j0}$ in the DB to the vector request **V** could be defined in accordance with the rule:

$$d_E(\mathbf{V}, \mathbf{V}_{j0}) = \ \min, \ \text{if} \ \left|\sum_{m=0}^{R-1} \mathrm{v(m)} \times \mathrm{v_{j_0}(m)}\right| = \max \quad \text{for} \ \mathrm{j_0} = 1, 2, .., \mathrm{J}. \tag{4.34}$$

Additional acceleration for the calculations in Eq. 4.34 is achieved, when the following suggestions are taken into account:

- In case, that $v(m) \geq 0$and $v_j(m) \geq 0$ the following relation is satisfied:

  $$\sum_{m=0}^{R-1} \mathrm{v(m)} \times \mathrm{v_j(m)} < \mathrm{v_{max}} \sum_{m=0}^{R-1} \mathrm{v_j(m)}\,, \tag{4.35}$$

  where:

  $$\mathrm{v_{max}} = \max\{v(m)\} \ \text{for m} = 0, 1, .., \ \mathrm{R} - 1. \tag{4.36}$$

  Then, for the so-called “modified” squared Euclidean distance $D_E(\mathbf{V}_i, \mathbf{V}_j)$ (MEUD) follows:

  $$D_E(\mathbf{V}, \mathbf{V}_j) = \ \|\mathbf{V}\|^2 + \left\|\mathbf{V}_\mathrm{j}\right\|^2 - 2v_{\mathrm{max}} \sum_{m=0}^{R-1} \mathrm{v_j(m)} \leq d_E(\mathbf{V}, \mathbf{V}_j). \tag{4.37}$$

Then Eq. 4.34 could be transformed as follows:

$$d_E(\mathbf{V}, \mathbf{V}_{j0}) = \ \min, \ \text{if} \ \left[\mathrm{v_{max}} \sum_{m=0}^{R-1} \mathrm{v_{j0}(m)}\right] = \max \quad \text{for} \ \mathrm{j_0} = 1, 2, .., \mathrm{J}. \tag{4.38}$$

- In case, that the components $v(m)$ and $v_j(m)$ have positive and negative values, they should be transformed in such a way, that to have positive values only:

  $$\mathrm{v'(m)} = v(m) + \min\{v(m), v_j(m)\}, \tag{4.39}$$

  $$\mathrm{v_j} \ = v_j + \min\{v(m), v_j(m)\} \ \text{for m} = 1, \ 2, .., \mathrm{R} - 1. \tag{4.40}$$

In this case, for the modified distance $D_E(\mathbf{V},\mathbf{V}_j)$ is obtained:

$$D_E(\mathbf{V},\mathbf{V}_j)=\|\mathbf{V}\|^2+\|\mathbf{V}_\mathrm{j}\|^2-2v'_{\max}\sum_{m=0}^{R-1}\mathrm{v}'_\mathrm{j}(\mathrm{m})\le d_E(\mathbf{V},\mathbf{V}_j), \tag{4.41}$$

and the rule, represented by Eq. 4.38 is changed accordingly:

$$d_E(\mathbf{V},\mathbf{V}_{j_0})=\min,\quad \text{if}\ \left[\mathrm{v}'_{\max}\sum_{m=0}^{n-1}\mathrm{v}'_{\mathrm{j}_0}(\mathrm{m})\right]=\max\quad \text{for}\ \mathrm{j}_0=1,2,..,\mathrm{J}. \tag{4.42}$$

where $\mathrm{v}'_{\max}=\max\{v'(m)\}$ for $\mathrm{m}=0,1,..,\mathrm{R}-1$

The rule for the classification of the vector $\mathbf{V}$ (Eq. 4.30), based on the K-NN and MEUD is transformed as follows:

$$\mathbf{V}\in C_{\beta_0},\ \ if\ \ \mathrm{S}_1\left(\mathbf{V},\mathbf{V}^{\beta_1}_{\alpha_{\beta_1}}\right)\ge \mathrm{S}_2\left(\mathbf{V},\mathbf{V}^{\beta_2}_{\alpha_{\beta_2}}\right)\ge\ldots.\ge \mathrm{S}_\mathrm{K}\left(\mathbf{V},\mathbf{V}^{\beta_K}_{\alpha_{\beta_K}}\right), \tag{4.43}$$

where

$$\beta_0=\max\{\mathrm{h}(\beta_k)\}\ \text{for}\ \beta_k=1,2,..,\ \mathrm{Q}\ \text{and}\ \mathrm{k}=1,2,..,\mathrm{K}; \tag{4.44}$$

$$\mathrm{S}_\mathrm{k}\left(\mathbf{V},\ \mathbf{V}^{\beta_k}_{\alpha_{\beta_k}}\right)=\left[\mathrm{v}'_{\max}\sum_{m=0}^{R-1}\mathrm{v}'^{\beta_\mathrm{k}}_{\alpha_{\beta_\mathrm{k}}}(\mathrm{m})\right] \tag{4.45}$$

Comparing Eqs. 4.30, 4.32, and 4.33 with Eqs. 4.43, 4.44, and 4.45 respectively follows, that the use of the similarity criterion $\mathrm{S}_\mathrm{k}\left(\mathbf{V},\mathbf{V}^{\beta_k}_{\alpha_{\beta_k}}\right)$ instead of $\mathrm{d}_\mathrm{k}\left(\mathbf{V},\mathbf{V}^{\beta_k}_{\alpha_{\beta_k}}\right)$ does not influence the classification results for the vector $\mathbf{V}$, but the number of multiplications is reduced (n × J) times, with retained number of sums. Besides, the corresponding sum $\sum_{m=0}^{R-1}\mathrm{v}'^{\beta_\mathrm{k}}_{\alpha_{\beta_\mathrm{k}}}(\mathrm{m})$ for each vector $\mathbf{V}^{\beta_k}_{\alpha_{\beta_k}}$ could be defined in advance and saved in the DB metadata. In result, the computational complexity is significantly reduced and the classification of the vector-request $\mathbf{V}$-enhanced.

In order to improve the similarity arrangement for the first K-nearest vectors to the vector-request, in Eq. 4.43 is possible to use the criterion "Cosine similarity" (CSim) [10, 17], instead of $\mathrm{S}_k(.)$. In correspondence to this criterion, the nearness between the two vectors, $\mathbf{V}$ and $\mathbf{V}_j$ could be evaluated using the relation:

$$CSim(\mathbf{V},\mathbf{V}_j)=\frac{\sum_{m=0}^{R-1}\mathrm{v}'(\mathrm{m})\times\mathrm{v}'_\mathrm{j}(\mathrm{m})}{\left[\sqrt{\sum_{m=0}^{R-1}[\mathrm{v}'(\mathrm{m})]^2}\right]\times\left[\sqrt{\sum_{m=0}^{R-1}[\mathrm{v}'_\mathrm{j}(\mathrm{m})]^2}\right]} \tag{4.46}$$

In order to avoid “square root” operations, Eq. 4.46 could be transformed as Squared Cosine Similarity (SCSim) = (CSim)^2. Further reduction of the needed calculations is achieved using the Modified Squared Cosine Similarity (MSCSim), defined as follows:

$$MSCSim(\mathbf{V},\mathbf{V}_j) = \frac{\left[v'_{\max} \times \sum_{m=0}^{R-1} \mathrm{v}'_{\mathrm{j}}(\mathrm{m})\right]^2}{\left[\sum_{m=0}^{R-1} [\mathrm{v}'(\mathrm{m})]^2\right] \times \left[\sum_{m=0}^{R-1} [\mathrm{v}'_{\mathrm{j}}(\mathrm{m})]^2\right]} \tag{4.47}$$

Then, in Eq. 4.43 the term $\mathrm{S}_k(.)$ is substituted by MSCSim(.). In result, the number of multiplications, needed for the calculation of MSCSim is reduced n times, and the search of the closest vector in a DB, containing J vectors, to the vector request, is enhanced (n.J) times. The efficiency of the presented approach for enhanced search of closest vector grows up with the increasing of the number of vectors in the DB.

Experimental Results

For the experiments was used the software implementation of the method in C + +, Windows environment. Significant part of the experiments aimed to prove the efficiency of the Modified Mellin-Fourier Transform. For the experiments was used the well-known test image “Lena”, 256 × 256 pixels, 8 bpp. The experiments were performed for various values of the main parameters: the side of the subscribed circle, the number of discrete radiuses, etc.

On Fig. 4.5a is shown the experimental image “Lena”, on which are indicated the points, which participate in the EPT (the black points are not retained and the image is restored after corresponding interpolation). On Fig. 4.5b is shown the discretization grid of LPT, and on Fig. 4.5c – the points, which participate in the LPT (the used points here are marked as black). This experiment confirms the efficiency of the new approach, because in the well-known LPT the retained central part of the processed image is smaller.

Some of the experimental results performed with the same test image “Lena”, are given below on Figs. 4.6 and 4.7.

The next part of the experiments was aimed at the content-based object retrieval. For this were used three specially developed image databases of the Technical University of Sofia: the first contained 180 faces of adult people, the second – 200 faces of adult people and children and the third – more than 200 scanned documents. Most of the faces in the databases are cropped from larger images. These photos were taken in various lighting conditions with many shadows, different views, etc. Very good results were obtained for search of similar faces in the databases. In the test database of adults, were included the images “Lena”, rotated in 90° and 270° and scaled up cropped part of the same original test image.

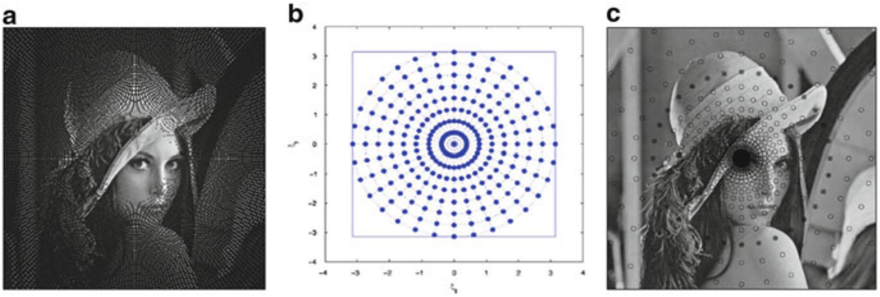


Fig. 4.5 Experimental images which indicate the points, retained after EPT and LPT. (**a**) Points retained by EPT. (**b**) Discretization grid of LPT. (**c**) Points retained by LPT

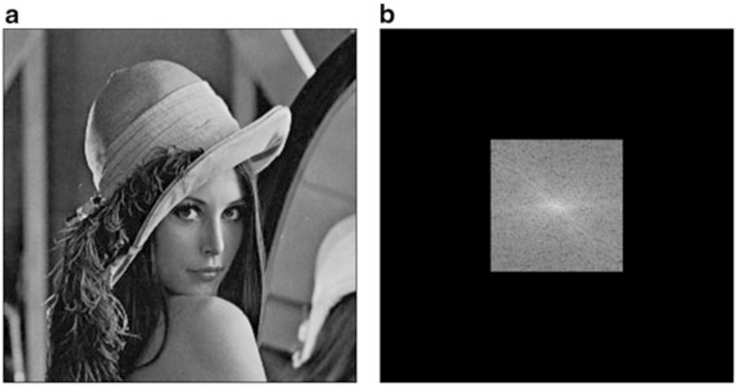


Fig. 4.6 Original test image and the result after DFT when the window of selected coefficients is of size H = 96. (**a**) Original image "Lena". (**b**) After first DFT

The image request (one of the cropped images from the test image "Lena") was the upper-left one in Fig. 4.8. The experiments proved the method efficiency. The results obtained confirm the RST-invariance of the method representation: the first five images are of the test image "Lena" – the first two are scaled up and cropped; the next three are the original, and the same image rotated on 180° and 90° correspondingly.

The experiments aimed at the detection of children sexual abuse in multimedia files need additional pre-processing. For this, color image segmentation was first performed, in order to detect naked parts of human bodies and then these parts were extracted from the images and defined as individual objects. After that the object search in the corresponding database was initiated. Special attention was paid to ability for children and adults faces recognition. The experiments confirm that this recognition is successful enough.

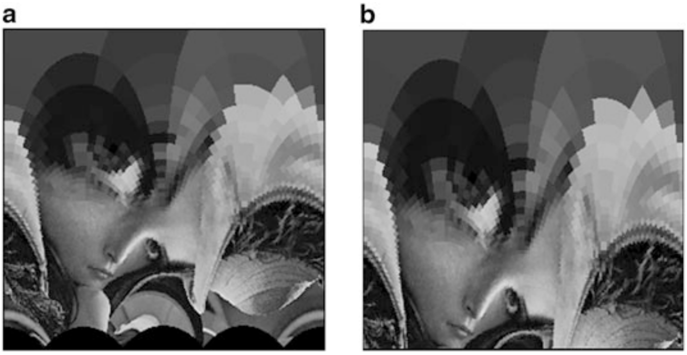
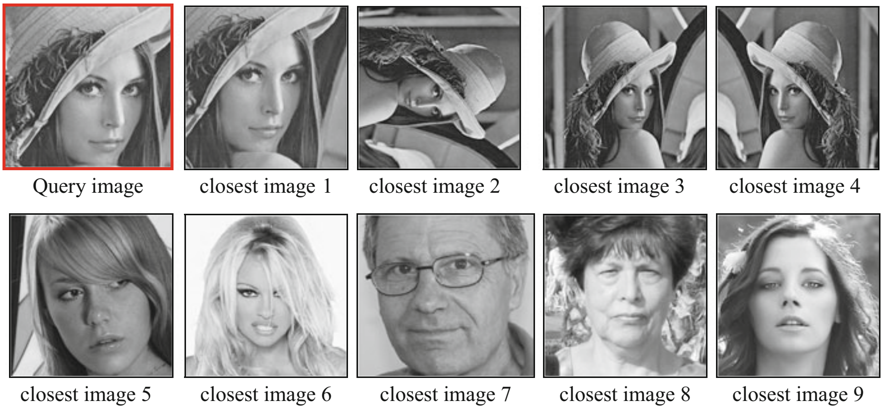


Fig. 4.7 Results after EPT for two values of the window of selected coefficients. (**a**) The test image after EPT, for H = 96. (**b**) The test image after EPT, for H = 64



Query image closest image 1 closest image 2 closest image 3 closest image 4

closest image 5 closest image 6 closest image 7 closest image 8 closest image 9

Fig. 4.8 The first K = 9 closest images from a database of 180 faces

On Fig. 4.9 are shown some of the results obtained for search of child's face in a mixed database, containing 200 faces of children and adults. Each image in the database was classified as belonging to one of these two classes. The experiments were performed under following conditions: $B_{max} = 255$, $K = 11$ and $H = 128$. In accordance with Eq. 4.18 the size $R = 888$ of RSTC vector is calculated for $\alpha = 20$.

The images on Fig. 4.9 are the closest to the image request (upper-left) in the test database. The experiment confirmed the method reliability when the searched face is of same person: the images 1, 2, and 4 in the first row are of the same child as the image request. The situation is same with images 6, 7, and 10 in spite of the fact that in the database comprised photos of more than 40 children. The error in this search result is one face only – the image 8. In some cases is possible to get large number of wrong images in the selection. In order to solve possible uncertainties, the final decision is taken in correspondence with Eq. 4.28 – i.e., with longer vector.

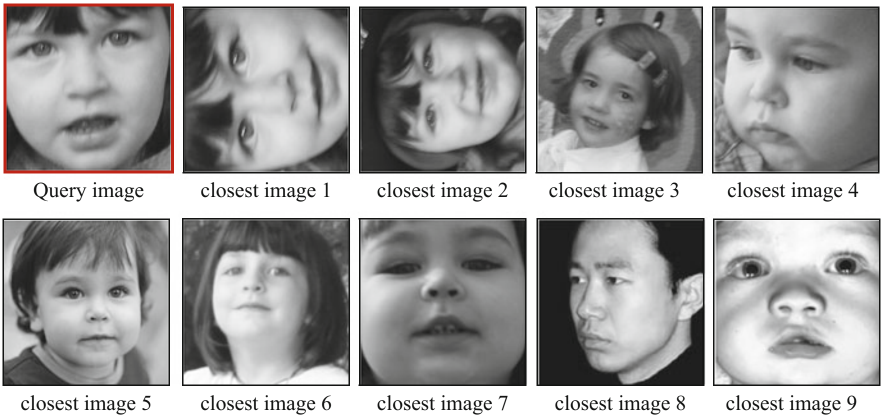

Query image
closest image 1
closest image 2
closest image 3
closest image 4
closest image 5
closest image 6
closest image 7
closest image 8
closest image 9

Fig. 4.9 Results for first $K = 9$ closest images obtained to image request (*upper left*) in a mixed image database of 200 faces (children and adults)

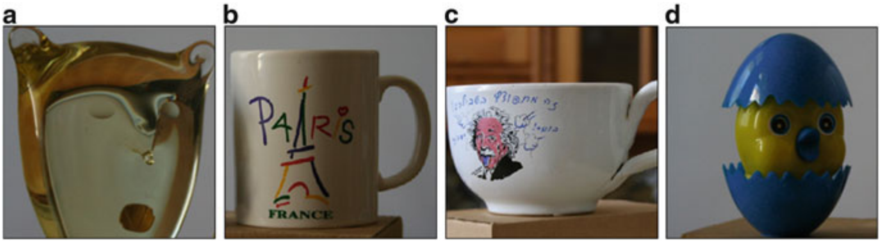


Fig. 4.10 Various 3D objects (**a**) Glass; (**b**) Mug; (**c**) Cup; (**d**) Egg

Significant attention was paid for the detection of 3D objects, represented by several multi-view 2D images. For this, was used the special database of the Technical University of Sofia, containing more than 200 multi-view 2D images of various 3D objects. All images were of size 256 × 256 pixels; 24 bpp. Each object was represented by four views, placed at equal distances in a sector of 20°. Another set of four images was taken from positions, placed at 5° up. For illustration, here is used a small database of four objects (Fig. 4.10), represented by 30 images. Each object had four views. More views (4.8) were used in the database for the object "Cup" only. Each image represented the object in different scale.

The experiments proved that the 3D object recognition is reliable in a view angle of 20°. On Fig. 4.11 are shown the results obtained for one of the test images. The image request ("Mug") is at the upper left corner of Fig. 4.11, and after it on the same row follow the closest images detected in the database (the names of the images are in accordance with their number in the database). The remaining closest images are arranged in the next two rows.

The graphic representation of the distances, calculated for the corresponding vectors of the objects represented by the 2D images in the database, is shown on

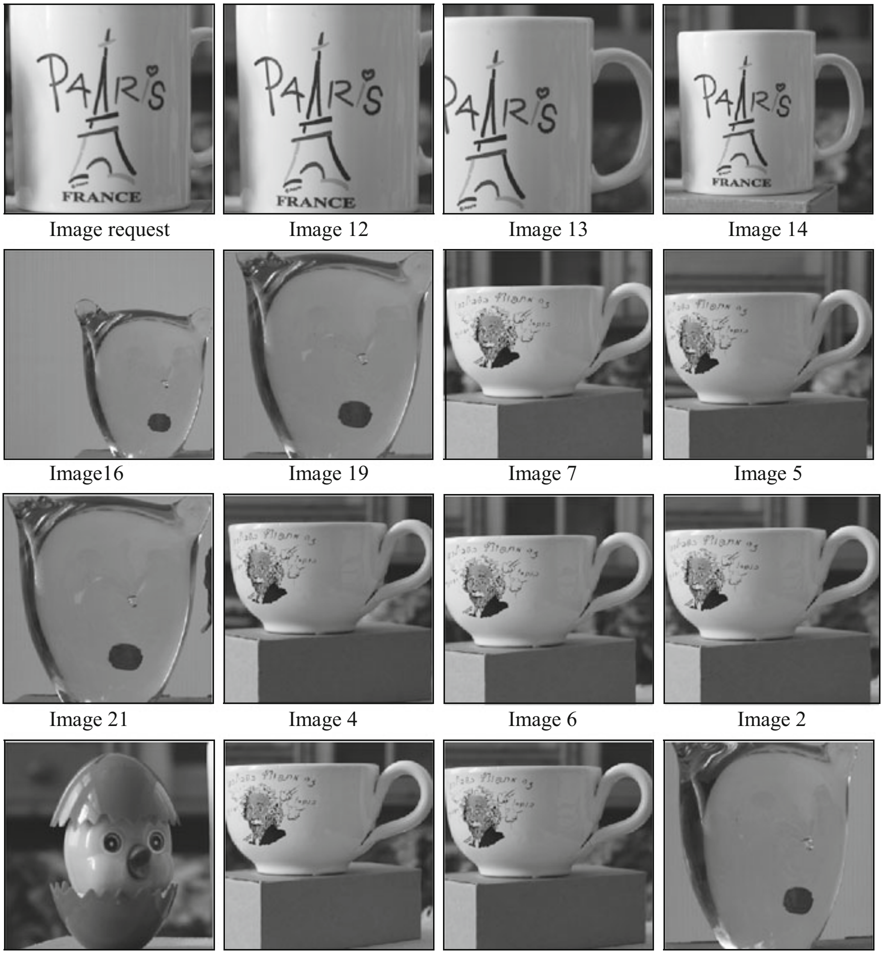


Image request Image 12 Image 13 Image 14

Image16 Image 19 Image 7 Image 5

Image 21 Image 4 Image 6 Image 2

Image 9 Image 1 Image 3 Image 22

Fig. 4.11 Results for the image request in a database of multi-view images

Fig. 4.12. It is easy to notice, that the first closest images (No. 1–3, corresponding to images 12–14 from Fig. 4.11) are of the same object. Besides, they are the only images of this object in the test database. The vectors of next closest images are at much larger distance, which proves their belonging to another object. This example was used for illustration purposes only. In real application tasks, special decision rules should be set, developed in accordance with the objects features.

Another group of experiments was aimed at the analysis of scanned documents. The database of scanned documents, comprised images of scanned texts, and signatures ($Q = 2$). The database contained more than 100 samples ($P > 100$) of each class; texts comprised examples of Latin and Cyrillic alphabets, printed

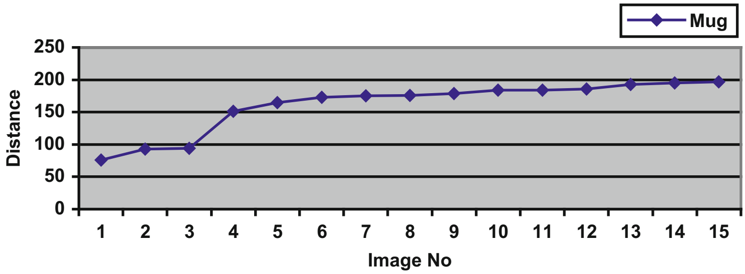


Fig. 4.12 Distances between the vector of the image request and the vectors of closest objects in the database

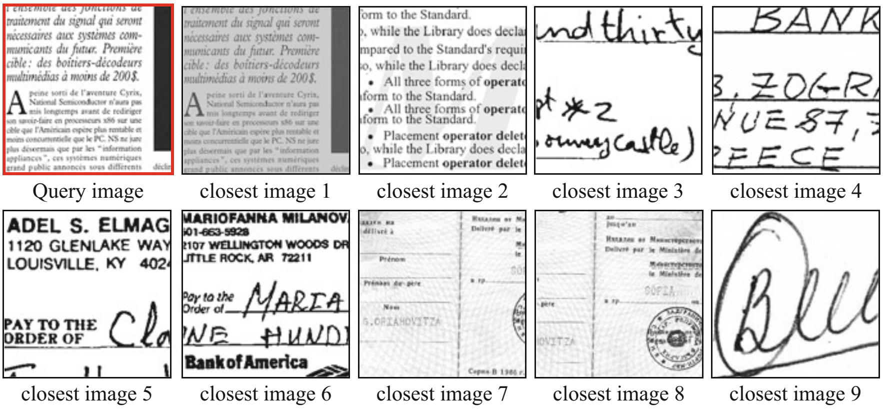
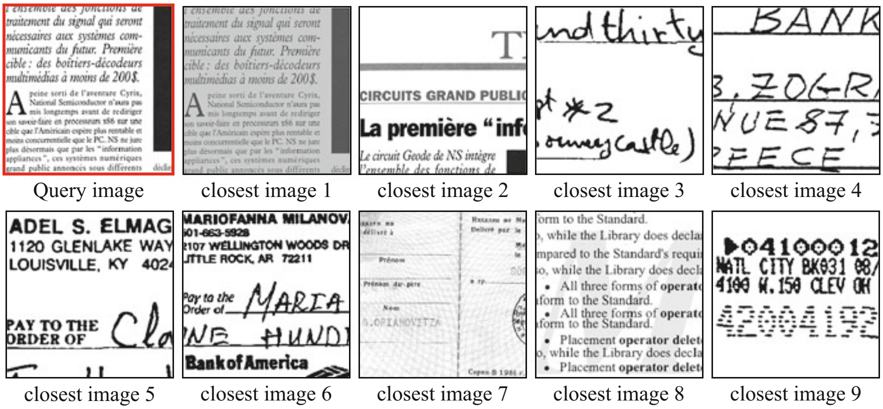

Query image
closest image 1
closest image 2
closest image 3
closest image 4
closest image 5
closest image 6
closest image 7
closest image 8
closest image 9

Fig. 4.13 Results obtained for image request in a database with 100 scanned texts documents for $H = 96$ and $K = 94$

and handwritten texts. All images were of size 256 × 256 pixels, grayscale (8 bpp, n = 256); $B_{max} = 255$. The experiments were performed for two versions of vector generation: *1st version*: the size of the retained coefficients square (H) equal to 96; *2nd version*: the size of the retained coefficients square (H) equal to 128. In Figs. 4.13 and 4.14 are shown results obtained for one of the test images (the image request is the upper left one) and the closest K = 11 images. In most cases (90 %) the information provided for H = 96 was enough for the right classification, but for some test images we had small number of mistakes. In Fig. 4.13 is given the result for version 1. There are three mistakes, i.e. the last three images (signatures) were classified as text, instead as belonging to the class of signatures. As it is seen in Fig. 4.14, the use of version 2 (H = 128) ensured the right decision (there are no mistakes at all). One of the images in the DB (the second in Figs. 4.13 and 4.14) was the same as the image request, but with changed contrast. In both experiments it was qualified as closest, which proves the method invariance to contrast changes.



Query image

closest image 1

closest image 2

closest image 3

closest image 4

closest image 5

closest image 6

closest image 7

closest image 8

closest image 9

Fig. 4.14 Results obtained for image request in a database with 100 scanned texts documents for $H = 128$ and $K = 9$

Conclusions

In this paper is presented a method for invariant object representation with Modified MFT based on the IPD. The main differences from the famous MFT are: (a) the first DFT is performed for limited number of coefficients only. In result is obtained an approximated image representation, suitable for the object representation; (b) instead of the Log-Pol Transform, here was used the Exponential-Polar Transform (EPT), in accordance with the description in Step 7 of the algorithm. As a result, the part of the participating points from the matrix of the Fourier coefficients' modules is larger (i.e. bigger central part of the image participates in the EPT and correspondingly – in the object description). The number of coefficients, used for the object representation is additionally limited in accordance with the vector length selection. Besides, the new transform is invariant to contrast changes because of the normalization performed in Step 10.

In result, the MMFT described above has the following advantages over the MFT:

- The number of transform coefficients used for the object representation is significantly reduced and from this naturally follows the lower computational complexity of the method, which permits real-time applications.
- The choice of coefficients, used for the vector calculation offers wide possibilities by setting large number of parameters, each of relatively wide range, which permits the method use in various applications: content- and context-based object retrieval in image databases, face recognition, etc.
- A new, simplified algorithm for fast search of closest vector in large image databases is developed.
- The method permits to detect similarity of 3D objects, represented by multiple 2D views, in a view angle of 200. Additional experiments show that the view angle could be wider for 3D objects of high symmetry.

The new approach, presented in this work, permits reliable object detection and identification in various positions, lighting conditions and view points.

Acknowledgements This paper was supported by the System Research and Application (SRA) Contract No. 0619069. This work was also supported in part by the Joint Research Project Bulgaria-Romania (2010–2012): "Electronic Health Records for the Next Generation Medical Decision Support in Bulgarian and Romanian National Healthcare Systems".

References

1. Nixon M, Aguado A (2002) Feature extraction and image processing. Newness, Oxford
2. Costa L, Cesar R (2001) Shape analysis and classification: theory and practice. CRC, LLC, Boca Raton
3. Lowe D (2005) Distinctive image features from scale-invariant keypoints. Int J Comput Vis 60(2):91–110
4. Carneiro G, Jepson A (2004) Flexible spatial models for grouping local image features. CVPR 2:747–754
5. Theodoridis S, Koutroumbas K (2009) Pattern recognition, 4th edn. Academic, Amsterdam/Boston/Heidelberg/London/New York/Oxford/Paris/San Diego/San Francosko/Singapore/Sidney/Tokyo
6. Schmid C, Dorko G, Lazebnik S, Mikolajczyk K, Ponce J (2005) Pattern recognition with local invariant features. In: Chen CH, Wang PS-P (eds) Handbook of pattern recognition and computer vision. World Scientific, Singapore/Hackensack, pp. 71–92
7. Cootes T, Taylor C, Cooper D (1995) Active shape models-their training and application. CVIU 61(1):38–59
8. Lu Z, Li D, Burkhardt H (2006) Image retrieval based on RST-invariant features. IJCSNS 6(2A):169–174
9. Reddy B, Chatterji B (1996) An FFT-based technique for translation, rotation, and scale-invariant image registration. IEEE Trans Image Process 5(8):1266–1271
10. Javidi B (ed) (2002) Image recognition and classification: algorithms, systems and applications. Marcel Dekker, New York
11. Derrode S, Ghorbel F (2001) Robust and efficient Fourier-Mellin transform approximations for gray-level image reconstruction and complete invariant description. Comput Vis Image Underst 83(1/1):57–78. Elsevier
12. Kountchev R, Rubin S, Milanova M, Todorov Vl, Kountcheva R (2009) Non-linear image representation based on IDP with NN. WSEAS Trans Signal Process 9(5):315–325
13. Kountchev R, Todorov Vl, Kountcheva R (2010) RSCT-invariant object representation with modified Mellin-Fourier transform. WSEAS Trans Signal Process 6(4):196–207
14. Kountchev R, Rubin S, Milanova M, Kountcheva R (2011) Invariant object description with inverse pyramid based on the truncated modified discrete Mellin-Fourier transform. In: The 2011 IEEE international conference on information reuse and integration (IEEE IRI'11), Las Vegas, pp 360–365
15. Caltech database "Faces". http://www.vision.caltech.edu/Image_Datasets/faces/. Accessed 26 June 2013
16. Wu K, Lin J (2000) Fast VQ encoding by an efficient kick-out condition. IEEE Trans Circuits Syst Video Technol 10(1):59–62
17. Qian G, Sural S, Gu Y, Pramanik S (2004) Similarity between euclidean and cosine angle distance for nearest neighbor queries. In: Proceedings of the 2004 ACM symposium on applied computing, Nicosia, pp 1232–1237