Feature extraction for CBIR and Biometrics applications

RYSZARD S. CHORAS
University of Technology & Life Sciences
Faculty of Telecommunications & Electrical Engineering
S. Kaliskiego 7, 85-796 Bydgoszcz
POLAND

Abstract: In CBIR (Content-Based Image Retrieval), visual features such as shape, color and texture are extracted to characterize images. Each of the features is represented using one or more feature descriptors. During the retrieval, features and descriptors of the query are compared to those of the images in the database in order to rank each indexed image according to its distance to the query. In biometrics systems images used as patterns (e.g. fingerprint, iris, hand etc.) are also represented by feature vectors. The candidates patterns are then retrieved from database by comparing the distance of their feature vectors. The feature extraction methods for this applications are discussed.

Key–Words: CBIR, Biometrics, Feature extraction

1 Introduction

A CBIR (Content-Based Image Retrieval) and Biometrics systems can be viewed as two main components: feature extraction and the search engine. These systems are based on computer vision methods to solve the image retrieval problem; i.e., the problem of searching for images in large image databases.
Color, texture, local shape and spatial information, in a variety of forms, are the most widely used features in such systems. In response to a user's query, the system returns images that are similar in some user-defined sense.

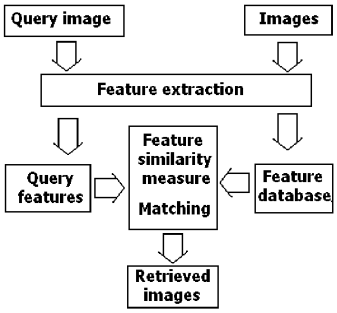


Figure 1: Schematic diagram of the image retrieval process.

The basic idea of the CBIR is to compactly describe an image by a feature vector and then match query images to the most resemblant image within the database according to the similarity of their features (Fig. 1).
CBIR can be divided in the following steps:

Preprocessing: The image is first processed in order to extract the features, which describe its contents. The processing involves filtering, normalization, segmentation, and object identification. The output of this stage is a set of significant regions and objects.

Feature extraction: Features such as shape, texture, color, etc. are used to describe the content of the image. Image features can be classified into primitives.

A biometric system is a pattern recognition system that recognizes a person on the basis of a feature vector derived from a specific physiological or behavioral characteristic that the person possesses.
Invariant features are extracted from the signal for representation purposes in the feature extraction subsystem. During the enrollment process, a representation (called template) of the biometrics in terms of these features is stored in the system. The matching subsystem accepts query and reference templates and returns the degree of match or mismatch as a score, i.e., a similarity measure. A final decision step compares the score to a decision threshold to deem the comparison a match or non-match. The personal attributes used in a biometric identification system can be physiological, such as facial features, fingerprints, iris, retinal scans, hand and finger geometry; or behavioral, the traits idiosyncratic of the individual, such as voice print, gait, signature, and keystroking.

A generalized diagram of a biometric system is shown in Figure 2. The component which is of great importance is the feature extraction algorithm. Feature extraction algorithm produces a feature vector, in which the components are numerical characterizations of the underlying biometrics.

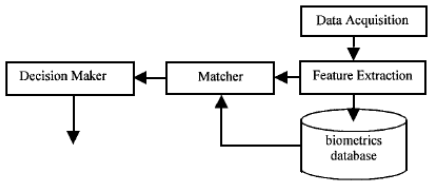


Figure 2: A generic biometrics-based system.

Let us represent an $n \times n$ color image by the vector function I such that $I(x, y)$ denotes the (r, g, b) vector at location x, y and records how red, green and blue a pixel appears. Suppose, we have a database M of m images: $M = I_1, I_2, \ldots, I_m$ and a query image I_q. Image indexing is all about finding the subset of images Q in M which are close to I_q. Mathematically, we might write:

$$Q = I_i : ||I_i - I_q||_d < T, \quad I_i \in M \quad (i = 1, 2, \ldots, m) \tag{1}$$

where $||.||_d$ is a distance measure which quantifies the similarity of two images and Q contains all those images in M which are sufficiently similar (their distance is below some user defined threshold T, to the query image I_q).

2 Representation of image content

Images features are divided to primitive and semantic features. Primitive features are those features that relate to the physical appearance of the image. Among them we can list:
- aspect ratio of the image;
- file format;
- color depth: black and white, n-bit grayscale, n-bit color;
- color: average color, color histogram or color correlation for the image or a subset of its pixels;
- texture: physical features of a part (or all) of the image when considered as a single texture;
- edge information: orientation, position and length of edges detected in the image or a subset of it;
- shapes: contour, orientation, elongation, size, bounding rectangle of shapes in the image;
- regions: areas of the image corresponding to homogeneous areas of the image;

Semantic features are abstract representations of images at different levels of detail, corresponding to human perception of the images.

As we mentioned before features should be extracted automatically from the images. Automatic extraction can be used only for the most primitive features, like color (computing the average color, the color histogram or color covariances of an area of the image) or size of a region of the image.

2.1 Color

Image characterized by color features have many advantages:

- Robustness. The color histogram is invariant to rotation of the image on the view axis, and changes in small steps when rotated otherwise or scaled [15]. It is also insensitive to changes in image and histogram resolution and occlusion.

- Effectiveness. There is high percentage of relevance between the query image and the extracted matching images.

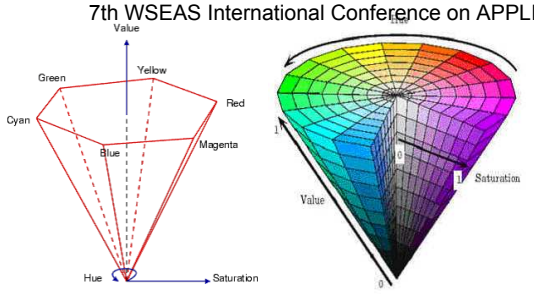
- Implementation simplicity. The construction of the color histogram is a straightforward process, including scanning the image, assigning color values to the resolution of the histogram, and building the histogram using color components as indices.

- Computational simplicity. The histogram computation has $O(M^2)$ complexity for images of size $M \times M$. The complexity for a single image match is linear, $O(n)$, where n represents the number of different colors, or resolution of the histogram.

- Low storage requirements. The color histogram size is significantly smaller than the image itself, assuming color quantisation.

As interesting color space we consider the Hue-Saturation-Value *(HSV)* color space, because it is compatible to the human color perception. Hue *(H)* is represented as angle. The purity of colors is defined by the saturation *(S)*, which varies from 0 to 1. The darkness of a color is specified by the value component *(V)*, which varies also from 0 (root) to 1 (top level).
The coordinate system and the *HSV* color model are shown in Figure 3 (a) and (b).

In case *RGB* to *HSV* conversion, the obtainable *HSV* colors lie within a triangle whose vertices are defined by the three primary colors in *RGB* space.



a) b)

Figure 3: a) *HSV* coordinate system b) *HSV* color model

The conversion formula is as follows :

$$H = cos^{-1}\left\{\frac{\frac{1}{2}[(R-G)+(R-B)}{\sqrt{(R-G)^2+(R-B)(G-B)}}\right\}$$
$$S = 1 - \frac{3}{R+G+B}[min(R,G,B)] \quad (2)$$
$$V = \frac{1}{3}(R+G+B)$$

Color features used in image retrieval include global and local color histograms, the mean (i.e., average color), and higher order moments of the histogram [18]. Average and dominant colors can be used to filter out irrelevant images without too much computational cost. However, they do not support a detailed comparison of the color appearance among images. The global color histogram provides a good approach to the retrieval of images that are similar in overall color content.
An image histogram refers to the probability mass function of the image intensities. This is extended for color images to capture the joint probabilities of the intensities of the three color channels. More formally, the color histogram is defined by,

$$h_{A,B,C}(a,b,c) = NProb(A=a, B=b, C=c) \quad (3)$$

where A, B and C represent the three color channels (*R,G,B* or *H,S,V*) and N is the number of pixels in the image.

A color image $I(x,y)$ of size $X \times Y$, which consists of three channels $I = (I_R, I_G, I_B)$, the color histogram used here is

$$h_c(m) = \frac{1}{XY}\sum_{x=0}^{X-1}\sum_{y=0}^{Y-1}\begin{cases} 1 & if \quad I(x,y)\ in\ bin\ m \\ 0 & otherwise \end{cases} \quad (4)$$

where a color bin is defined as a region of colors.

Let h and g represent two color histograms. The euclidean distance between the color histograms h and g can be computed as:

$$d^2(h,g) = \sum_A\sum_B\sum_C(h\sum(a,b,c) - g(a,b,c))^2 \quad (5)$$

The intersection of histograms h and g is given by:

$$d(h,g) = \frac{\sum_A\sum_B\sum_C \min(h(a,b,c), g(a,b,c))}{\min(|h|,|g|)} \quad (6)$$

where $|h|$ and $|g|$ gives the magnitude of each histogram, which is equal to the number of samples.

For *HSV* space distance formula is given by

$$d_{ij} = 1 - \frac{1}{\sqrt{5}}[(v_i - v_j)^2 + (s_i\cosh_i - s_j\cosh_j)^2 + (s_i\sinh_i - s_j\sinh_j)^2]^{1/2} \quad (7)$$

which corresponds to the proximity in the *HSV* color space.

Color moments have been successfully used in many retrieval systems especially when the image contains just the object. The first order (mean), the second (variance) and the third order (skewness) color moments have been proved to be efficient and effective in representing color distributions of images. Mathematically, the first three moments are defined

$$\mu_c = \frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N} f_c(x,y) \quad (8)$$

$$\sigma_c = (\frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}(f_c(x,y) - \mu_c)^2)^{\frac{1}{2}} \quad (9)$$

$$s_c = (\frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}(f_c(x,y) - \mu_c)^3)^{\frac{1}{3}} \quad (10)$$

where $f_c(x,y)$ is the value of the c-th color component of the image pixel (x,y), and MN is the number of pixels in the image.

Since only 9 (three moments for each of the three color components) numbers are used to represent the color content of each image, color moments are a very compact representation compared to other color features.The similarity function used for retrieval is a weighted sum of the absolute differences between the suitable moments.

2.2 Texture

Texture is a powerful regional descriptor that helps in the retrieval process. Texture, on its own does not have the capability of finding similar images, but it

can be used to classify textured images from non-textured ones and then be combined with another visual attribute like color to make the retrieval more effective.

Statistical methods, including Fourier power spectra, co-occurrence matrices, shift-invariant principal component analysis (SPCA), Tamura features, Wold decomposition, Markov random field, fractal model, and multi-resolution filtering techniques such as Gabor and wavelet transform, characterize texture by the statistical distribution of the image intensity.

The co-occurrence matrix $C(i,j)$ counts the co-occurrence of pixels with gray values i and j at a given distance d. The distance d is defined in polar coordinates (d,α), with discrete length and orientation. In practice, α takes the values $0°$; $45°$; $90°$; $135°$; $180°$; $225°$; $270°$; and $315°$. The co-occurrence matrix $C(i,j)$ can now be defined as follows:

$$C(i,j) = Pr(I(p_1) = i \wedge I(p_2) = j \mid |p_1 - p_2| = d) \tag{11}$$

where Pr is probability, and p_1 and p_2 are positions in the gray-scale image I.

Texture features which can be extracted from gray level co-occurrence matrices are as follows:
Angular Second Moments

$$\sum_i \sum_j C(i,j)^2 \tag{12}$$

Correlation

$$\frac{\sum_i \sum_j (ij) C(i,j) - \mu_i \mu_j}{\sigma_i \sigma_j} \tag{13}$$

Variance

$$\sum_i \sum_j (i-j)^2 C(i,j) \tag{14}$$

Inverse Difference Moment

$$\sum_i \sum_j \frac{1}{1+(i-j)^2} C(i,j) \tag{15}$$

Entropy

$$-\sum_i \sum_j C(i,j) log C(i,j) \tag{16}$$

Inertia (or contrast)

$$\sum_i \sum_j (i-j)^2 C(i,j) \tag{17}$$

Cluster Shade

$$\sum_i \sum_j ((i-\mu_i) + (j-\mu_j))^3 C(i,j) \tag{18}$$

Gabor filters have been successfully applied in various computer vision applications and to texture analysis and image retrieval. The general functionality of the 2D Gabor filter family can be represented as a Gaussian function modulated by a complex sinusoidal signal. Specially, a 2D Gabor filter $g(x,y)$ can be formulated as

$$g(x,y;F,\theta) = \frac{1}{2\pi\sigma_x\sigma_y} \exp[-\frac{1}{2}(\frac{\bar{x}^2}{\sigma_x^2} + \frac{\bar{y}^2}{\sigma_y^2})] \exp[2\pi j F \bar{x}] \tag{19}$$

here

$$\begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}, \quad j = \sqrt{-1}$$

and

- σ_x and σ_y are the scaling parameters of the filter and determine the effective size of the neighborhood of a pixel in which the weighted summation (convolution) takes place,
- θ $(\theta \in [0,\pi])$ specifies the orientation of the Gabor filters,
- F is the radial frequency of the sinusoid.

Gabor filters worked as local bandpass filters and each filter is fully determined by choosing the four parameters $\{\theta, F, \sigma_x, \sigma_y\}$. Assuming that N filters are needed in an application, $4N$ parameters need to be optimized. The orientation parameter θ should satisfy $\theta \in [0,\pi)$. W is the radial frequency of the Gabor filter and is application dependent. σ_x and σ_y are the effective sizes of the Gaussian functions and are within the range $[\sigma_{min}, \sigma_{max}]$.

The Gabor filter $g(x,y;F,\theta)$ forms complex valued function. Decomposing $g(x,y;F,\theta)$ into real and imaginery parts gives

$$g(x,y;F,\theta) = r(x,y;F,\theta) + ji(x,y;F,\theta) \tag{20}$$

where

$$\begin{aligned} r(x,y;F,\theta) &= g(x,y;F,\theta)\cos(2\pi F,\bar{x}) \\ i(x,y;F,\theta) &= g(x,y;F,\theta)\sin(2\pi F,\bar{x}) \end{aligned} \tag{21}$$

The Gabor filtered output of an image $I(x,y)$ is obtained by the convolution of the image with the Gabor function $g(x,y;F,\theta)$. Given a neighborhood window of size $W \times W$ for $W = 2t+1$, the discrete convolutions of $I(x,y)$ with respective real and imaginery components of $g(x,y;F,\theta)$ are

$$C_{ev}(x,y;F,\theta) = \sum_{l=-t}^{t} \sum_{m=-t}^{t} I(x+l, y+m) r(x,y;F,\theta) \tag{22}$$

$$C_{odd}(x,y;F,\theta) = \sum_{l=-t}^{t}\sum_{m=-t}^{t} I(x+l,y+m)i(x,y;F,\theta) \quad (23)$$

The channel output is computed as

$$C(x,y;F,\theta) = \sqrt{(C_{ev}(x,y;F,\theta))^2 + (C_{odd}(x,y;F,\theta))^2} \quad (24)$$

After applying Gabor filters on the image with different scale s and orientation k we obtain an array of magnitudes. These magnitudes represent the energy content at different scale and orientation of the image (Figure 4). The following mean μ_{sk} and

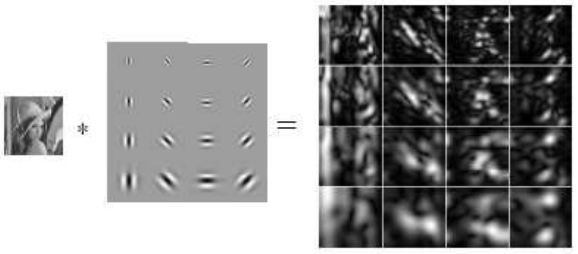


Figure 4: Gabor filters convolution

standard deviation S_{sk} of the magnitude of the transformed coefficients are used to represent the homogenous texture feature of the region

$$\mu_{sk} = \frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N} C_{sk}(x,y;F,\theta) \quad (25)$$

$$S_{sk} = \sqrt{\sum_{x=1}^{M}\sum_{y=1}^{N}(C_{sk}(x,y;F,\theta) - \mu_{sk})^2} \quad (26)$$

where $s = 0,1,\ldots,S-1$ and $k = 0,\ldots,K-1$.

The feature vector *(FV)* is constructed using μ_{sk} and S_{sk} as feature components.

2.3 Shape

Shape based image retrieval is the measuring of similarity between shapes represented by their features. Shape content description is difficult to define because measuring the similarity between shapes is difficult. Therefore, two steps are essential in shape based image retrieval, they are: feature extraction and similarity measurement between the extracted features. Shape descriptors can be divided into two main categories: region-based and contour-based methods. Region-based methods use the whole area of an object for shape description, while contour-based methods use only the information present in the contour of an object. The shape descriptors described here are:

- shape descriptors - features calculated from objects contour: circularity, aspect ratio, discontinuity angle irregularity, length irregularity, complexity, right-angleness, sharpness, directedness.Those are translation, rotation (except angle), and scale invariant shape descriptors.

It is possible to extract image contours from the detected edges. We extract and store a set of shape features from the contour image and for each individual contour. These features are (Figure 5):

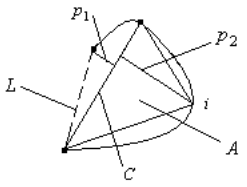


Figure 5: Shape and measures used to compute features.

1. Circularity $cir = \frac{4pA}{P^2}$.
2. Aspect Ratio $ar = \frac{p_1+p_2}{C}$.
3. Discontinuity Angle Irregularity
   $dar = \sqrt{\frac{(\sum|\theta_i-\theta_{i+1}|)}{2\pi(n-2)}}$.
   A normalized measure of the average absolute difference between the discontinuity angles of polygon segments made with its adjoining segments.
4. Length Irregularity
   $lir = \frac{\sum|L_i-L_{i+1}|}{K}$,
   where $K = 2P$ for $n > 3$ and $K = P$ for $n = 3$.
5. Complexity $com = 10^{\frac{-3}{n}}$. A measure of the number of segments in a boundary group weighted such that small changes in the number of segments have more effect in low complexity shapes than in high complexity shapes.
6. Right-Angleness $ra = \frac{r}{n}$. A measure of the proportion discontinuity angles which are approximately right-angled.
7. Sharpness
   $sh = \sum\frac{max(0,1-(\frac{2|\theta-\pi|}{\pi})^2)}{n}$.
   A measure of the proportion of sharp discontinuities (over 90).
8. Directedness $dir = \frac{M}{\sum P_i}$. A measure of the proportion of straight-line segments parallel to the mode segment direction.

where n - number of sides of polygon enclosed by segment boundary, A - area of polygon enclosed by segment boundary, P - perimeter of polygon enclosed by segment boundary, C - length of longest boundary chord, p_1, p_2 - greatest perpendicular distances from longest chord to boundary, in each half-space either side of line through longest chord, θ_i - discontinuity angle between $(i-1)$th and ith boundary segment, r - number of discontinuity angles equal to a right-angle within a specified tolerance, and M - total length of straight-line segments parallel to mode direction of straight-line segments within a specified tolerance.

- region-based shape descriptor utilizes a set of Zernike moments calculated within a disk centered at the center of the image.

Zernike moment of order n and repetition m is defined as:

$$Z_{nm} = \frac{n+1}{\pi} \iint_{x^2+y^2 \leq 1} V_{nm}(\rho, \theta) f(x, y) dx dy \quad (27)$$

where:
- $f(x, y)$ is the image intensity at (x, y) in Cartesian coordinates,
- $V_{nm}(\rho, \theta)$ is a complex conjugate of $V_{nm}(\rho, \theta) = R_{nm}(\rho) e^{-jm\theta}$ in polar coordinates (ρ, θ) and $j = \sqrt{-1}$,
- $n \geq 0$, and $n - |m|$ is even positive integer.

The polar coordinates (ρ, θ) in the image domain are related to the Cartesian coordinates (x, y) as $x = \rho cos(\theta)$ and $y = \rho sin(\theta)$.
$R_{nm}(\rho)$ is a radial defined as follows:

$$R_{nm}(\rho) = \sum_{s=0}^{\frac{n-m}{2}} \frac{(-1)^s [(n-s)!] \rho^{n-2s}}{s!(\frac{n+|m|}{2} - s)!(\frac{n-|m|}{2} - s)!} \quad (28)$$

The first six orthogonal radial polynomials are:

$$\begin{array}{ll} R_{00}(\rho) = 1 & R_{11}(\rho) = \rho \\ R_{20}(\rho) = 2\rho^2 - 1 & R_{22}(\rho) = \rho^2 \\ R_{31}(\rho) = 3\rho^3 - 2\rho & R_{33}(\rho) = \rho^3 \end{array} \quad (29)$$

The discrete approximation of Equation (27) is given as:

$$Z_{nm} = \frac{4(n+1)}{(N-1)^2 \pi} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} f(k, l) R_{nm}(\rho_{k,l}) e^{-jm\theta_{kl}} \quad (30)$$

$$0 \leq \rho_{k,l} \leq 1$$

where the discrete polar coordinates:

$$\rho_{k,l} = \sqrt{x_k^2 + y_l^2} \quad ; \quad \theta_{kl} = arctan(\frac{y_l}{x_k}) \quad (31)$$

are transformed by:

$$x_k = \frac{\sqrt{2}}{N-1} k + \frac{-1}{\sqrt{2}} \quad ; \quad y_l = \frac{\sqrt{2}}{N-1} l + \frac{-1}{\sqrt{2}} \quad (32)$$

for $k = 0, \ldots, N-1$ and $l = 0, \ldots, N-1$.

To calculate the Zernike moments of an image $f(x, y)$, the image is first mapped onto the unit disk using polar coordinates, where the center of the image is the origin of the unit disk. Pixels falling outside the unit disk are not used in the calculation.

Because Z_{mn} is complex, we use the Zernike moments modules $|Z_{mn}|$ as the features of shape in the recognition of patterns.

3 Applications

The CBIR technology has been used in several applications such as fingerprint identification, biodiversity information systems, crime prevention, medicine, among others. Some of these applications are presented in this section

3.1 Medical applications

Queries based on image content descriptors can help the diagnostic process. Visual features can be used to find images of interest and to retrieve relevant information for a clinical case. One example is a content-based medical image retrieval that supports mammographical image retrieval.
The main aim of the diagnostic method in this case is to find the best features and get the high classification rate for microcalcification and mass detection in mammograms.

The microcalcifications are grouped into clusters based on their proximity. A set of the features was initially calculated for each cluster:

- Number of calcifications in a cluster
- Total calcification area / cluster area
- Average of calcification areas
- Standard deviation of calcification areas
- Average of calcification compactness
- Standard deviation of calcification compactness
- Average of calcification mean grey level

- Standard deviation of calcification mean grey level
- Average of calcification standard deviation of grey level
- Standard deviation of calcification standard deviation of grey level.

Mass detection in mammography is based on shape and texture based features. The features are listed below :

- Mass area. The mass area, $A = |R|$, where R is the set of pixels inside the region of mass, and $|.|$ is set cardinal.
- Mass perimeter length. The perimeter length P is the total length of the mass edge. The mass perimeter length was computed by finding the boundary of the mass, then counting the number of pixels around the boundary.
- Compactness. The compactness C is a measure of contour complexity versus enclosed area, defined as: $C = \frac{P^2}{4\pi A}$ where P and A are the mass perimeter and area respectively. A mass with a rough contour will have a higher compactness than a mass with smooth boundary.
- Normalized radial length. The normalized radial length is sum of the Euclidean distances from the mass center to each of the boundary coordinates, normalized by dividing by the maximum radial length.
- Minimum and maximum axis. The minimum axis of a mass is the smallest distance connecting one point along the border to another point on the border going through the center of the mass. The maximum axis of the mass is the largest distance connecting one point along the border to another point on the border going through the center of the mass.
- Average boundary roughness.
- Mean and standard deviation of the normalized radial length. The mean μ and standard deviation σ of the normalized radial length are computed as

$$\mu_i = \frac{1}{n}\sum_{k=1}^{n} R_k \tag{33}$$

$$\sigma = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(R_k - \mu_i)^2} \tag{34}$$

where R_k is the normalized radial length at boundary point (x_k, y_k).

- Eccentricity. The eccentricity characterizes the lengthiness of a ROI. To this purpose a symmetric matrix A is defined as follows

$$\begin{aligned} A_{11} &= \sum_{i=1}^{N}(x_i - X_0)^2 \\ A_{12} = A_{21} &= \sum_{i=1}^{N}(x_i - X_0)(y_i - Y_0) \\ A_{22} &= \sum_{i=1}^{N}(y_i - Y_0)^2 \end{aligned} \tag{35}$$

where N is the number of the ROI pixels; x_i and y_i are the coordinates of a generic pixel, X_0 and Y_0 are the coordinates of the geometric center of the ROI. If λ_1 and λ_2 are the eigenvalue of the A matrix, in the elliptical approximation of the ROI region, the semi-axis values will be

$$S_1 = \sqrt{|\frac{\lambda_1}{2}|} \quad S_2 = \sqrt{|\frac{\lambda_2}{2}|} \tag{36}$$

Then the eccentricity is given by

$$eccentricity = \frac{S_1}{S_2} \tag{37}$$

with $S1 < S2$. An eccentricity close to 1 denotes a ROI like a circle, while values close to zero mean more stretched ROIs.

- Roughness. The roughness index was calculated for each boundary segment (equal length) as

$$R(j) = \sum_{k=j}^{L+j} |R_k - R_{k+1}| \tag{38}$$

for $j = 1, 2, \ldots, \frac{n}{L}$ where $R(j)$ is the roughness index for the jth fixed length interval.

- Average mass boundary. The average mass boundary calculated as averaging the roughness index over the entire mass boundary

$$R_{ave} = \frac{L}{n}\sum_{j=1}^{\frac{L}{n}} R(j) \tag{39}$$

where n is the number of mass boundary points and L is the number of segments.

3.2 Iris recognition

A typical iris recognition system often includes iris capture, preprocessing, feature extraction and feature matching. In iris recognition algorithm, preprocessing and feature extraction are two key processes. Iris preprocessing, including localization, segmentation, normalization and enhancement, is a basic step in iris identification algorithm. Iris feature extraction is the most important step in iris recognition, which determines directly the value of iris characteristics in actual application. Typical iris recognition system is illustrated in Fig. 6.

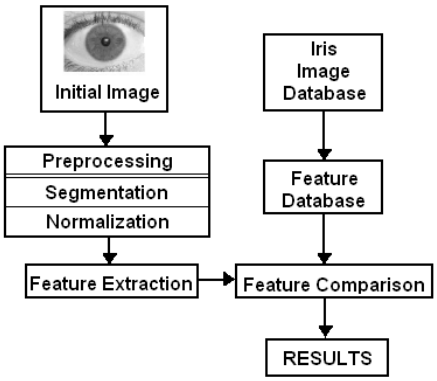


Figure 6: Typical iris recognition stages

Robust representations for iris recognition must be invariant to changes in the size, position and orientation of the patterns. Irises from different people may be captured in different sizes and, even for irises from the same eye, the size may change due to illumination variations and other factors. In order to compensate the varying size of the captured iris it is common to translate the segmented iris region, represented in the cartesian coordinate system, to a fixed length and dimensionless polar coordinate system. The next stage is the feature extraction.

The following formulas perform the transformation.

$$\theta \in [0, 2\pi], \quad \rho \in [0, 1], \quad I(x(\rho, \theta), y(\rho, \theta)) \rightarrow I(\rho, \theta) \tag{40}$$

$$\begin{aligned} x(\rho, \theta) &= (1 - \rho)x_p(\theta) + \rho x_i(\theta) \\ y(\rho, \theta) &= (1 - \rho)y_p(\theta) + \rho y_i(\theta) \end{aligned} \tag{41}$$

$$\begin{aligned} x_p(\theta) &= x_{p0}(\theta) + \rho_p \cos(\theta) \\ y_p(\theta) &= y_{p0}(\theta) + \rho_p \sin(\theta) \end{aligned} \tag{42}$$

$$\begin{aligned} x_i(\theta) &= x_{i0}(\theta) + \rho_i \cos(\theta) \\ y_i(\theta) &= y_{i0}(\theta) + \rho_i \sin(\theta) \end{aligned} \tag{43}$$

where $I(x, y)$ is the iris region, (x, y) and (ρ, θ) are the Cartesian and normalized polar coordinates respectively, (x_p, y_p) and (x_i, y_i) are coordinates on pupil and limbus boundaries along the θ direction, (x_{p0}, y_{p0}), (x_{i0}, y_{i0}) are the coordinates of pupil and iris centers.

The remapping is done so that the transformed image is rectangle with dimension 512×32 (Fig. 7).

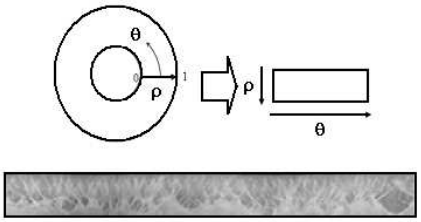


Figure 7: Transformed region

Most of iris recognition systems are based on Gabor functions analysis in order to extract iris image features. It consists in convolution of image with complex Gabor filters which is used to extract iris feature. As a product of this operation, complex coefficients are computed. In order to obtain iris signature, complex coefficients are evaluated and coded.

The normalized iris images (Fig. 7) are divided into two stripes, and each stripe into $K \times L$ blocks. The size of each block is $k \times l$. Localization of blocks is shown in Fig. 8.
Each block is filtered by

$$Gab(x, y, \alpha) = \sum_{x-\frac{k}{2}}^{x+\frac{k}{2}} \sum_{y-\frac{k}{2}}^{y+\frac{k}{2}} I(x, y) \cdot g(x, y) \tag{44}$$

The orientation angles of this set of Gabor filters are

$$\langle \alpha_i | \alpha_i = \frac{i\pi}{4}\ , \ \ i = 0, 1, 2, 3 \rangle \tag{45}$$

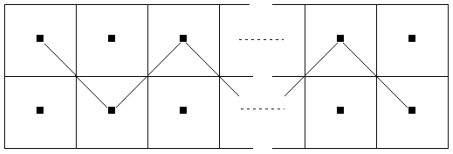


Figure 8: Localization of blocks.

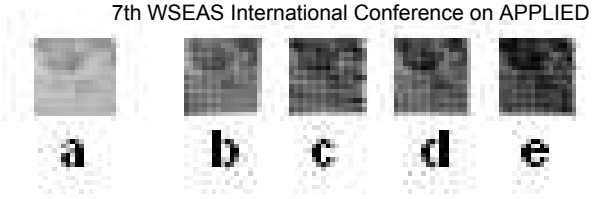


Figure 9: Original block iris image (a) and real part of $Gab(x, y, \alpha_i)$ for $\alpha_i = 0^\circ$ (b), $\alpha_i = 45^\circ$ (c), $\alpha_i = 90^\circ$ (d), $\alpha_i = 135^\circ$ (e)

To encode the iris we used the real part of (44) as

$$Code(x, y) = 1 \quad if \quad Re(Gab(x, y, \alpha_i) \geq th$$
$$Code(x, y) = 0 \quad if \quad Re(Gab(x, y, \alpha_i) < th \quad (46)$$

The iris binary *Code* can be stored as personal identify feature.

Figure 10: Iris Code

References:

[1] W.W. Boles and B. Boashash, A human identification technique using images of the iris and wavelet transform, *IEEE Transactions on Signal Processing*, vol. 46, no. 4, pp. 1185-1188, 1998.

[2] J.G. Daugman, High confidence visual recognition of persons by a test of statistical independence, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 11, pp. 1148-1161, 1993.

[3] J.G. Daugman, Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression, *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 11691179, 1988.

[4] D. Gabor, Theory of communication, *J. Inst. Elect. Eng.*, vol. 93, pp. 429459, 1946.

[5] A.K. Jain, R.M. Bolle, and S. Pankanti, Eds., *Biometrics: Personal Identification in Networked Society*, Norwell, MA: Kluwer, 1999.

[6] L. Ma, T. Tan, Y. Wang, and D. Zhang, Personal identification based on iris texture analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1519 1533, 2003.

[7] L. Ma, Y. Wang, and T. Tan, Iris recognition based on multichannel Gabor filtering, in *Proc. 5th Asian Conf. Computer Vision*, vol. I, 2002, pp. 279283.

[8] R.P. Wildes, J.C. Asmuth, G.L. Green, S.C. Hsu, R.J. Kolczynski, J.R. Matey, and S.E. McBride, A machine vision system for iris recognition, *Mach. Vision Applicat.*, vol. 9, pp. 18, 1996.

[9] J.K. Kim, H.W. Park, Statistical textural features for detection of microcalcifications in digitized mammograms, *IEEE Transactions on Medical Imaging*, (18), 1999, 231-238.

[10] L. Shen, R.M. Rangayyan, J.E.L. Desautels, (1994) Application of shape analysis to mammographic calcifications, *IEEE Trans. Medical Imaging*, 13, 1994, 263-274.

[11] H.D. Cheng, X. Cai, X. Chen, L. Hu, X. Lou, Computer-aided detection and classification of microcalcifications in mammograms: a survey, *Pattern Recognition*, 36, 2003, 2967-2991.

[12] E. Saber, A.M. Tekalp, Integration of color, edge and texture features for automatic region-based image annotation and retrieval. *Electronic Imaging* 7, pp. 684–700, 1998.

[13] C. Schmid, R. Mohr, (1997) Local grey value invariants for image retrieval. *IEEE Trans Pattern Anal Machine Intell* 19, pp. 530–534, 1997.

[14] IEEE Computer, special issue on Content Based Image Retrieval, 28, 9, 1995.

[15] W.Y. Kim, Y.S. Kim, A region-based shape descriptor using Zernike moments, *Signal Processing: Image Communications*, vol.16, pp.95-102, 2000.