

## INVARIANT OBJECT RECOGNITION USING RADON-BASED TRANSFORM\*

Tomasz ARODŹ

*Institute of Computer Science  
AGH University of Science and Technology  
al. Mickiewicza 30, 30-059 Kraków, Poland  
e-mail: arodz@icsr.agh.edu.pl*

Manuscript received 13 October 2004; revised 28 February 2005

Communicated by Steve Maybank

**Abstract.** The properties of the Radon transform are used to derive the transformation invariant to translation, rotation and scaling. The invariant transformation involves translation compensation, angle representation and 1-D Fourier transform. The new object recognition method is studied experimentally in two domains, mammogram labels recognition and face recognition. For mammogram labels, the recognition accuracy is 97 %, while in case of faces it reaches 96 %.

**Keywords:** Invariant object recognition, radon transform, face recognition, mammography

### 1 INTRODUCTION

One of the problems in the field of object recognition is to develop a method, which is invariant to certain variations in the input images. Such variations include translation, rotation-on-the-plane, scaling, as well as rotation-in-depth and illumination. In this paper, first three of the above-mentioned variations are studied. A transformation, which gives identical results if applied to a pair of images that are in the similarity relation with each other, is derived. The result of such a transformation can be treated as a new set of features for recognition, the set that is invariant to

---

\* This paper extends and modifies the ideas presented in a paper [1] at the ICCS '04 Conference

translation, rotation-on-the-plane and scaling of the input picture. The proposed invariant is suitable for recognition of 2D images. For 3D objects, some constraints on the variance of the input object have to be met, so that the object is pictured from the same camera-angle.

The methods for invariant object recognition can be divided into two groups. In the first, more straightforward approach, the value of the translation, scaling and rotation is explicitly estimated, by e.g. estimating the centroid and then analysing the shape of the object. Then, the object appearance in the image is normalised, e.g. by moving the centroid to the centre of the image and by explicit rescaling and rotating the object. This approach is carried out solely in the image domain. The drawback is that the estimation procedures may be susceptible to noise. The second group of methods does not involve estimation and normalisation steps. A transform is formulated, which preserves only that part of the information contained in the object appearance, which is invariant to the analysed variations. In such methods, the drawback is the inherent loss of some information, which may impair the recognition process.

In this paper, we focus on the second approach for achieving invariance. Several examples of such methods have already been studied. For example, the approaches based on geometric moments, such as Hu moments [5] and Zernike moments [16], are used. The Fourier transform is a basis for a group of methods, utilizing log-polar or Fourier-Mellin transforms [9, 13]. Other similar transforms include Hessian or Taylor invariants [3]. Methods based on shape of the object, i.e., not using its interior, are also employed for invariant recognition of objects. These include e.g. multi-vector eigenvector shape descriptors [8] and wavelet-based method [7]. A different approach using a set of random cross-sections through the object have also been proposed [15]. Finally, a method based on higher-order spectra and Radon transform has been formulated in [10]. In this paper, the Radon transform is considered as a basis for developing a new invariant transform. Similarly to [10], the proposed method starts with calculation of the Radon transform of the input image. However, we use translation compensation, angular representation and Fourier transform, whereas in [10], a method based on bispectral analysis of the Radon transform followed by centroid normalisation of the spectra is utilised.

The rest of the paper is arranged in the following way. Section 2 presents the Radon transform and its properties. In Section 3, the invariant method based on this transform is derived. Next, Section 4 describes the setup for experiments using mammography labels taken from the mini-MIAS dataset [12], and Yale [2] and BioID [6] faces datasets. Section 5 presents and discusses the results of the experiments. Finally, Section 6 summarizes the conclusions of the paper.

## 2 RADON TRANSFORM AND ITS PROPERTIES

The Radon transform of the translated, rotated and scaled images exhibits interesting properties, which can be employed to construct a method for invariant object

recognition. Therefore, the behaviour of the transform for these three variations in the input image should be defined.

For an image  $f : \mathbb{R} \times \mathbb{R} \rightarrow [0, 1]$  containing an object (i.e. not uniformly black), the result  $g$  of the Radon transform  $RAD$  is a function  $g : \mathbb{R} \times [0, 2\pi) \rightarrow \mathbb{R}_+$  defined as:

$$g(s, \theta) = RAD(f(x, y)) = \int_{-\infty}^{\infty} f(s \cos \theta - u \sin \theta, s \sin \theta + u \cos \theta) du \quad (1)$$

$$\begin{bmatrix} s \\ u \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (2)$$

Assuming that the images contain an object to be recognized, for each  $\theta$  a non-zero  $g(s, \theta)$  can be found. As shown in [10], any translation in spatial domain leads in the Radon domain to translation in the  $s$  direction. The amount of the translation varies with the  $\theta$  dimension. The scaling of the original image along both axes results in the scaling along the  $s$  axis in the Radon domain. The value of the transform is also scaled. The rotation in spatial domain leads to circular translation along the  $\theta$  axis in the Radon domain. The behaviour of the Radon transform is summarized in Table 1 and depicted in Figure 1.

Image	Input image function $f$	Resulting Radon transform $g = RAD(f)$
Original	$f(x, y)$	$g(s, \theta)$
Translated	$f(x - x_0, y - y_0)$	$g(s - x_0 \cos \theta - y_0 \sin \theta, \theta)$
Scaled	$f(\alpha x, \alpha y)$	$\frac{1}{ \alpha } g(\alpha s, \theta)$
Rotated	$f_{\text{polar}}(r, \phi + \theta_0)$	$g(s, (\theta + \theta_0) \bmod 2\pi)$

Table 1. Behaviour of the Radon transform for translated, scaled and rotated images

The Radon transform of a continuous function is also continuous. The transform of the function with bounded support, such as a real-world image, has bounded support. Therefore, for a finite set of pictures, the coordinates can be adjusted, so that for each picture  $f(x, y)$ ,  $\text{supp}(RAD(f)) \subset [0, 1) \times [0, 2\pi)$ .

### 3 RADON-BASED INVARIANT RECOGNITION

To derive the translation, rotation and scale invariance from the original Radon transform, a set of simple steps is introduced. To simplify the notation, in places where it will not lead to confusion,  $g(s)$  will denote the values of  $g(s, \theta)$  for any specified  $\theta$ . Moreover, in places where the relation between original and translated or scaled image is studied, the functions and variables associated with the latter image will be capitalized. The subscripts  $T$ ,  $S$  and  $R$  will signify that the function is invariant to translation, scaling and rotation, respectively.

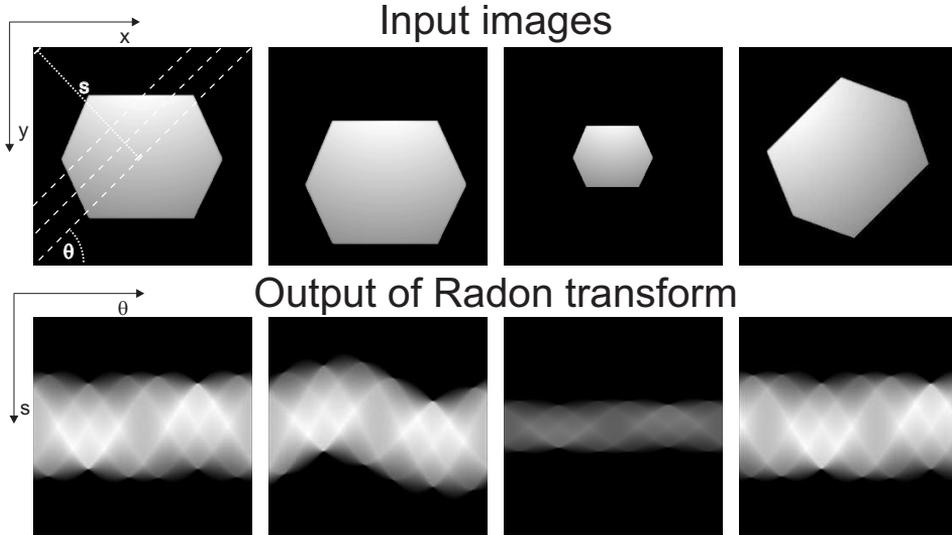


Fig. 1. The Radon transform and its behaviour for translated, scaled and rotated input images

### 3.1 Invariance to Translation

Direct translation of the Radon transform result along the  $s$  axis can lead to the translation invariance, when the value of the translation,  $s_g(\theta)$ , is chosen according to the following formula:

$$s_g = \inf \{s : g(s, \theta) > 0\}, \tag{3}$$

for a given  $\theta$ . For different  $\theta$ , the values of  $s_g$  may be different.

As the following theorem shows, the function  $g_T(s) = g(s + s_g)$  is translation invariant and preserves the scale variance.

**Theorem 1.** Let  $g(s)$  and  $G(\alpha s + s_0)$  be two functions that are two versions of each other, translated by  $s_0$  and scaled by  $\alpha$ , i.e.,  $\alpha g(s) = G(\alpha s + s_0)$ . The functions  $g_T(s)$  and  $G_T(\alpha s)$  are scaled version of each other, thus  $g_T(s)$  is translation invariant and preserves the scaling.

**Proof.** For functions  $g(s)$  and  $G(\alpha s + s_0)$ , the values of the direct translation  $s_g$  and  $s_G$ , defined by (3), are related by the equation  $s_G = \alpha s_g + s_0$ . Thus, the function  $g_T$  and  $G_T$  are scaled version of each other, i.e.,  $\alpha g_T(s) = G_T(\alpha s)$ :

$$\alpha g_T(s) = \alpha g(s + s_g) = G(\alpha(s + s_g) + s_0) = G(\alpha s + s_G) = G_T(\alpha s), \tag{4}$$

and the transformation from  $g(s)$  to  $g_T(s)$  is indeed translation invariant and has a property of preserving the function scaling.  $\square$

For  $g_T(s)$ , the relation:  $\text{supp}(g_T) \subset [0, 1) \times [0, 2\pi)$  holds. Moreover,  $g_T(0) > 0$ . The process of achieving translation invariance is depicted in Figure 2.

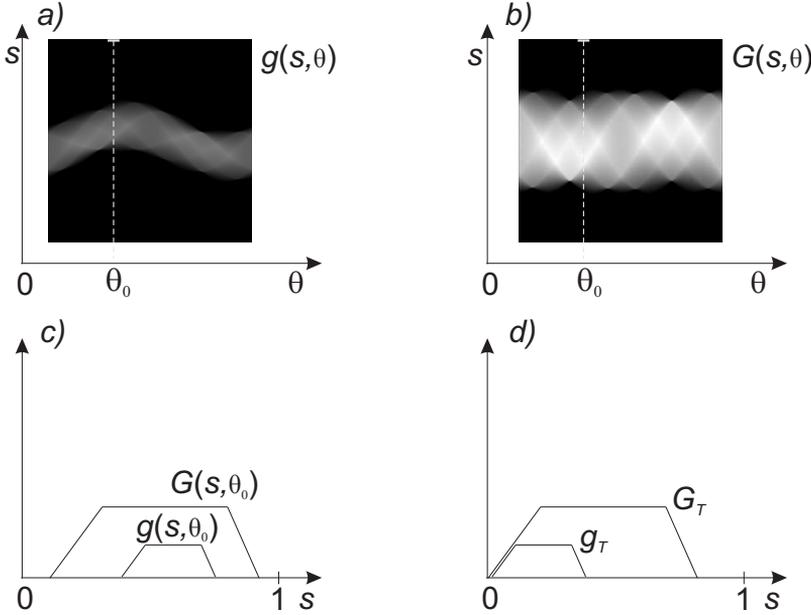


Fig. 2. The scheme demonstrating the translation invariance. (a–b) Two scaled and translated functions:  $g$  and  $G$  are considered. (c) Values of  $g$  and  $G$  for fixed  $\theta_0$ . (d) Translation invariant and scale preserving functions  $g_T$  and  $G_T$ . See text for details.

The translation invariant in this form results in an intermediate set of features, which is invariant to translation. This step uses the estimation of the boundary of the support of the image in the Radon domain. This boundary is related to the boundary of the object in the original image. This makes the invariant dependent on prior segmentation of the object from the image.

One should note that other well-defined points along the  $s$  axis could be used instead of  $s_g(\theta)$ , e.g. the value of  $s$  for which  $g$  is maximal for a given  $\theta$ . More advanced methods for achieving translation invariance could also be considered, e.g. the modulus of the Fourier transform along the  $s$  axis. However, such an approach does not preserve correctly the scale in discrete images. Another approach, using the formula relating the translation along the  $s$  dimension to a given  $\theta$  (see relations in Table 1), would require the estimation of the translation values  $x_0$  and  $y_0$  in the original image.

### 3.2 Invariance to Scaling

Within this section, it is assumed that the translation has been removed. As in the previous section, the result of the Radon transform is treated individually for each value of  $\theta$ . The derivation of scale invariance is based on the auxiliary function  $h(x)$  having the following properties.

**Lemma 1.** For a function  $g_T(s)$  with continuous first derivative in  $[0, 1]$ , the function  $h(x)$  defined as:

$$g_{\text{rev}}(s) = g_T(1-s), \quad \text{where } s \in [0, 1]; \quad (5)$$

$$h_{\text{rev}}(x) = \int_0^x \left| \frac{dg_{\text{rev}}}{ds} \right| ds, \quad \text{where } x \in [0, 1]; \quad (6)$$

$$h(x) = h_{\text{rev}}(1-x) \quad (7)$$

is well-defined, differentiable, nonincreasing and nonnegative in  $[0, 1]$ . Moreover,  $h(1) = 0$  and  $h(0) > 0$ .

**Proof.** From the assumptions on  $g_T(s)$ , the derivatives  $g'_T$  and also  $g'_{\text{rev}}$  are continuous. Thus, the integral in (6) is well-defined and  $h_{\text{rev}}(x)$  is continuous and differentiable. Since the integration is done for the absolute values,  $h_{\text{rev}}(x)$  is non-decreasing and nonnegative. Also,  $h_{\text{rev}}(0) = 0$ . The properties of  $h(x)$  stem from the properties of  $h_{\text{rev}}(x)$ , i.e.,  $h(x)$  is also nonnegative and continuous. Since  $h(x)$  is reversed version of  $h_{\text{rev}}(x)$ , it is nonincreasing and  $h(1) = 0$ . Furthermore, since  $\exists s \in [0, 1] : g_T(s) > 0$  and  $g_T(s)$  is continuous in  $[0, 1]$ , the integral over the whole range  $[0, 1]$  is nonzero, i.e.,  $h_{\text{rev}}(1) > 0$  and thus  $h(0) > 0$ .  $\square$

The function  $h(x)$  can be used to derive scale invariant transforms. First, let  $h_P(x)$  be a variant of the function  $h(x)$  restricted to the domain  $(0, P]$ , such that  $h([0, P]) > 0$  and  $h([P, 1]) = 0$ . Such a  $P$  can always be found in  $[0, 1]$ , since  $h(0) > 0$ ,  $h(1) = 0$  and  $h(x)$  is nonincreasing, according to Lemma 1.

**Lemma 2.** Let two mappings  $s_h^1$  and  $s_h^2$  be defined as:

$$s_h^1(\gamma) = \sqrt{x^2 + h_P(x)^2} \quad \gamma = \arctan\left(\frac{h_P(x)}{x}\right); \quad (8)$$

$$s_h^2(\gamma) = |h'_P(x)| \quad \gamma = \arctan\left(\frac{h_P(x)}{x}\right). \quad (9)$$

Both  $s_h^1(\gamma)$  and  $s_h^2(\gamma)$  are functions:  $s_h^1 : [0, \frac{\pi}{2}) \rightarrow \mathbb{R}_+$  and  $s_h^2 : [0, \frac{\pi}{2}) \rightarrow \mathbb{R}_+$ .

**Proof.** As  $h(x)$  and  $h_P(x)$  are nonincreasing and  $h_P((0, P]) > 0$ , the function  $\frac{h_P(x)}{x}$  is nonnegative, decreasing and continuous in its domain. Also,  $h_P(P) = 0$  and  $\lim_{x \rightarrow 0} \frac{h_P(x)}{x} = \infty$  since  $h(0) > 0$ . The  $\arctan(y)$  is continuous and increasing in  $[0, \infty)$ , with  $\arctan([0, \infty)) \rightarrow [0, \frac{\pi}{2})$ . Therefore, the  $\gamma = \arctan\left(\frac{h_P(x)}{x}\right)$  is bounded

to period  $(\frac{\pi}{2}, 0]$  for  $x \in (0, P]$  and is continuous and decreasing. That is, the left-hand sides of the definitions (8) and (9) are well-defined.

As for the right-hand sides of the definitions (8) and (9),  $h_P(x)$  is differentiable for  $x \in (0, P]$ . Both the derivative and square root in equations (8) and (9) are well defined in this range.  $\square$

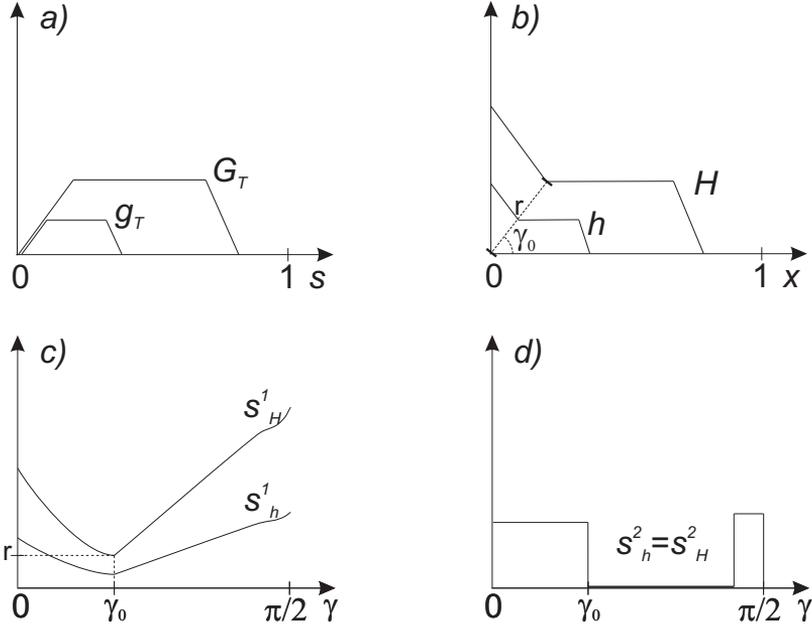


Fig. 3. The scheme demonstrating scale invariance. (a) Scaled functions  $g_T$  and  $G_T$ . (b) Auxiliary function  $h$ . (c) Scale invariant (after normalisation) functions  $s_h^1$  and  $s_H^1$ . (d) Scale invariant functions  $s_h^2$  and  $s_H^2$ . See text for details.

It can be shown that the function  $s_h^2(\gamma)$  is scale-invariant. The function  $s_h^1(\gamma)$ , after normalization, is also scale-invariant.

**Theorem 2.** Let  $g_T(s)$  and  $G_T(S)$  be two functions with the translation removed and meeting the constraints of Lemma 1, such that  $G_T(S) = \alpha g_T(s)$  and  $S = \alpha s$ , i.e., two functions that are scaled versions of each other. The function  $s_h^2(\gamma)$  and the function  $S_H^1(\gamma)$  defined as

$$S_H^1(\gamma) = \frac{s_h^1(\gamma)}{\int_0^{\frac{\pi}{2}} s_h^1(\varphi) d\varphi} \tag{10}$$

are both scale invariant, i.e.,  $s_H^2(\gamma) = s_h^2(\gamma)$  and  $S_H^1(\gamma) = s_h^1(\gamma)$ , where the functions  $h(x)$  and  $H(y)$  are defined on the basis of  $g_T(s)$  and  $G_T(S)$ , respectively, according to (7).

**Proof.** Since the function  $h_P(x)$  is derived from  $g_T(s)$  with only integration, differentiation, reorientation of the axis and restriction of the domain, the functions  $h_P(x)$  and  $H_P(y)$  are also scaled versions of each other, i.e.,  $H_P(y) = \alpha h_P(x)$  and  $y = \alpha x$ . Thus:

$$\gamma_H = \arctan\left(\frac{H_P(y)}{y}\right) = \arctan\left(\frac{\alpha h_P(x)}{\alpha x}\right) = \arctan\left(\frac{h_P(x)}{x}\right) = \gamma_h. \quad (11)$$

Also,

$$s_H^2(\gamma) = \frac{dH}{dy} = \frac{dH(\alpha x)}{d\alpha x} = \frac{1}{\alpha} \frac{d\alpha h(x)}{dx} = \frac{dh}{dx} = s_h^2(\gamma), \quad (12)$$

that is, the first of the functions is indeed scale invariant. Furthermore,

$$s_H^1(\gamma) = \sqrt{y^2 + H_P(y)^2} = \sqrt{(\alpha x)^2 + (\alpha h_P(x))^2} = \alpha \sqrt{x^2 + h_P(x)^2} = \alpha s_h^1(\gamma), \quad (13)$$

and thus

$$\int_0^{\frac{\pi}{2}} s_H^1(\gamma) d\gamma = \alpha \int_0^{\frac{\pi}{2}} s_h^1(\gamma) d\gamma. \quad (14)$$

Therefore,  $S_h^1(\gamma) = S_H^1(\gamma)$ , i.e., the normalized version of the second of the functions is also scale invariant.  $\square$

To simplify the notation, the functions  $s_h^2(\gamma)$  and  $S_h^1(\gamma)$  have been derived using one-dimensional function  $g_T(s)$  for fixed  $\theta$ , but, in fact, the functions are defined in two dimensions:  $s_h^2(\gamma, \theta)$  and  $S_h^1(\gamma, \theta)$  as  $g_T(s, \theta)$  is. In the following sections, the translation and scale invariant function  $g_{TS}(\gamma, \theta)$  will represent any of the above-defined two functions,  $S_h^1(\gamma, \theta)$  or  $s_h^2(\gamma, \theta)$ . The process of achieving scale invariance is depicted in Figure 3. While a more straightforward method for achieving invariance could be used, e.g. by estimating the value of the scaling from the length of the support of  $g_T$  for given  $\theta$  and then applying normalisation, our goal in this study is to minimize the steps relying on explicit variance estimation.

### 3.3 Invariance to Rotation

The rotation in the image is reduced to the circular translation in the  $\theta$  direction by the Radon transform and is preserved in this form by translation and scale variance elimination, both operating along the  $s$  axis. For a discrete approximation of  $g_{TS}(\gamma, \theta)$ , the magnitude of the 1-D Fourier transform along the  $\theta$  dimension

$$g_{TSR}(\gamma, \Theta) = |DFT(g_{TS}(\gamma, \theta))| \quad (15)$$

is rotation invariant, since the magnitude of the Fourier transform is invariant [16] with respect to circular translation of any function  $f(x)$  in  $[0, X]$ :

$$|DFT(f(x))| = |DFT(f((x + x_0) \bmod X))|. \quad (16)$$

While using only the magnitude discards information contained in the phase coefficients, the amount of information is still suitable for successful recognition. Indeed, such an approach is used in other methods for deriving e.g. translation invariance [9]. As the function  $g_{\text{TS}}(\gamma, \theta)$  is already translation and scale invariant,  $g_{\text{TSR}}(\gamma, \Theta)$  possesses the property of full invariance to similarity-preserving variations in the input image.

#### 4 EXPERIMENTAL SETUP

The developed invariant recognition method has been tested for two different types of objects: mammogram labels and faces. Mammogram labels (see Figure 5) are human-made, regular and homogenous within a given class. On the other hand, faces (see Figure 6) are natural, contain fine details and differ significantly even for the same person.

To test the method, the resulting transform  $g_{\text{TSR}}(\gamma, \Theta)$  of input image has to be compared with a set of labelled, transformed sample images. This has been done using the *k-nearest neighbour* method (for  $k = 3$ ). The *leave-one-out* technique [14] has been applied to obtain the classification accuracy. In turns, each of the images was recognised based on distances to the rest of the images for the given class and all images for other classes. The total number of turns thus equalled to the total number of images in the dataset. The number of correct classifications was divided by the total number of images to obtain the recognition accuracy. For choosing the nearest neighbours, three metrics for transformed images have been evaluated:

**Euclidean distance**  $d_E(x, y) = \sqrt{(x - y)^T (x - y)}$ ,

**Manhattan distance**  $d_M(x, y) = \sum_i |x_i - y_i|$ ,

**Tanimoto[14] dissimilarity measure**  $d_T(x, y) = 1 + \frac{(x-y)^T(x-y)}{x^T y}$ .

The invariant method has been implemented in the following way. The pre-processed images of size  $S \times S$  pixels were transformed using the Radon transform to obtain  $S\sqrt{2} \times S$  images. Thus, the sampling in the Radon transform along the  $\theta$  axis was set to  $\frac{2\pi}{S}$ . Then, translation invariance is achieved, followed by scale invariance. Both of the presented scale-invariance methods were tested. After the application of scale-invariance transform, the data points are spaced irregularly along the  $\gamma$  axis. This is due to the transformation from the  $x$  to the  $\gamma$  dimension involved in obtaining the scale invariance (8), (9). Therefore, piecewise cubic Hermite interpolation was used to obtain regular mesh of the size  $S \times S$  pixels. This type of interpolation preserves the shape and reproduces the monotonicity of the data. After interpolation, the rotation invariance is achieved, i.e., the magnitude of the FFT, of size  $S \times S$ , was calculated. Finally, it has been reshaped into a vector of  $S^2$  dimensions, which was treated as input to the nearest-neighbour scheme. The proposed methods were implemented in MATLAB, with values of  $S$  depending on the size of the input images, as specified in Section 4.1 and Section 4.2. The

result of the invariant for two rotated, scaled and translated images is presented in Figure 4.

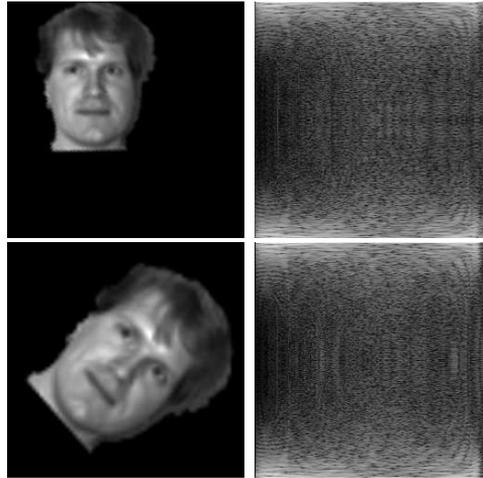


Fig. 4. Two original images (left) and logarithms of their representations after the application of the invariant transform (right)

#### 4.1 Invariant Mammogram Labels Recognition Experiments

The mammogram labels have been extracted from breast images from the mini-MIAS dataset [12]. They represent the information on the setup of the mammography (see Figure 5, left), i.e., depicted breast and exposition direction. In the dataset, two classes have been identified: representing the left breast (L) and the right breast (R). Both classes are captured in the medio-lateral (ML) projection (i.e., view from the centre of the chest outwards). Thus, the two classes will be further referred to as (L, ML) and (R, ML) for left and right breast, respectively.

Often, the label does not fit entirely in the image, and some of the labels are partially cropped (see Figure 5, right). Thus, two experiments were conducted, one taking into account only the whole labels (Figure 5, centre), and second using all labels. The dataset contains 66 labels, 17 whole and 22 cropped in class (L, ML), and 8 whole and 19 cropped in class (R, ML). Each label has been extracted from the mammogram and put at a random position within a  $256 \times 256$  pixel image with a uniform black background.

#### 4.2 Invariant Face Recognition Experiments

The data for the face experiments were taken from Yale faces dataset [2], consisting of 11 pictures for each of 15 individuals. These images are of relatively high resolution

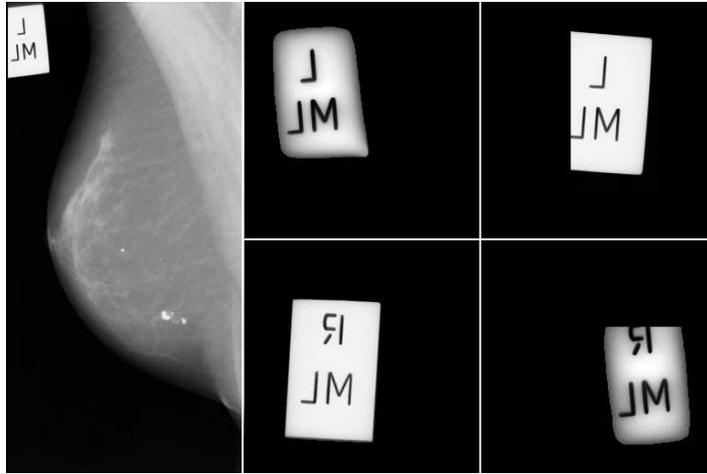


Fig. 5. Typical mammogram(left), mammogram labels (centre) and cropped mammogram labels (right) from mini-MIAS dataset [12]

( $256 \times 256$  pixels) and consist of the head and neck of a person. Additionally, in order to evaluate the performance of the method on images that contain only tightly cropped face region of lower resolution ( $32 \times 32$  pixels), the BioID dataset [6] consisting of 1520 images of 23 individuals was used. The images from both of the sets are depicted in Figure 6.



Fig. 6. Faces from the Yale[2] dataset (above) and BioID [6] dataset (below)

For the Yale dataset two pictures for each person, containing faces with left- and right-side illumination have been eliminated from the test sets. The original images have faced some pre-processing in order to be suitable for the tests. In original Yale images, the faces are pictured in a uniform white background. To be suitable for the proposed method the faces have been extracted from the original Yale images and put into images with uniform black background. Then all the images have been randomly rotated in the range of  $[-\frac{\pi}{2}, \frac{\pi}{2}]$  and randomly scaled in the range of [66 %, 100 %]. Thus, the largest image may be 1.5 larger than the smallest one. Finally, the objects were randomly translated within the  $256 \times 256$  picture.

In face recognition, the problem of the small number of available photos for each person is often encountered. To model this problem, a method alternative to the leave-one-out scheme was employed. For every person, the available images were split into the sample set and the testing set. Then, the nearest-neighbour method was used on the sample set for each image in the test set. For the Yale dataset, the sample set contained, for each individual, a single image labelled “normal”.

For the BioID dataset, 5 images for each of the 23 persons were chosen as a sample set. In several cases when the database did not contain 5 images of a particular person, the available number of images was used. As a test set, 200 images were randomly selected from the whole dataset. Since the images in the dataset consist of faces with some background, each face was extracted from the images using the position of the eyes. As a result, a cropped face image of size  $32 \times 32$  pixels was created for each image.

The images contain also other types of variances than spatial ones, e.g. illumination. Thus, in tests, simple 64-level histogram equalization is performed to compensate for minor lighting variances. Another variant of lighting compensation is applied for the test scenario with full Yale set, with side-illuminated images included. This scenario is used to ease the comparison with other face recognition methods. In this compensation method, first, the image is histogram equalized. Then, *bior3.7* wavelet decomposition is performed [4]. Image reconstructed from the sixth-level wavelet approximation is subtracted from histogram equalized original image. Finally, the resulting image is normalized into  $[0, 1]$  pixel intensity range.

## 5 EXPERIMENTAL RESULTS AND DISCUSSION

For mammogram labels, the results are presented in Table 2. The results for two experiments are given, one using only the whole labels and the second one using also cropped labels.

	Scale inv. func.: $S^1$			Scale inv. func.: $s^2$		
	Tanimoto	Eucldn	Manhttn	Tanimoto	Eucldn	Manhttn
No cropped labels	0.96	0.96	0.96	0.96	0.88	0.88
All labels	0.89	0.97	0.89	0.71	0.70	0.73

Table 2. Results for 3-Nearest Neighbour and leave-one-out scheme, mammogram labels

In face recognition domain, the results of the tests for the Yale dataset in the leave-one-out scheme are summarized in Table 3. Table 4 presents the results for the second scenario, with 1 sample and 8 testing images per person. Since the calculation of the Radon transform of the  $256 \times 256$  pixels image is time consuming, the method has also been applied to images downsized to  $64 \times 64$  and  $32 \times 32$  pixels.

	Scale inv. func.: $S^1$			Scale inv. func.: $s^2$		
	Tanimoto	Eucldn	Manhttn	Tanimoto	Eucldn	Manhttn
$256 \times 256$	0.882	0.882	0.911	0.941	0.933	0.919
$64 \times 64$	0.896	0.896	0.896	0.963	0.956	0.963
$32 \times 32$	0.874	0.874	0.874	0.919	0.933	0.933

Table 3. Results for 3-Nearest Neighbour and leave-one-out scheme, 9 images per person, Yale faces dataset

	Scale inv. func.: $S^1$			Scale inv. func.: $s^2$		
	Tanimoto	Eucldn	Manhttn	Tanimoto	Eucldn	Manhttn
$256 \times 256$	0.808	0.808	0.858	0.909	0.892	0.892
$64 \times 64$	0.775	0.775	0.817	0.900	0.867	0.850
$32 \times 32$	0.733	0.733	0.708	0.775	0.758	0.775

Table 4. Results for 1-Nearest-Neighbour, 8 test and 1 sample image per person, Yale faces dataset

For the test scenario with single sample image per person, the recognition accuracy significantly decreases for  $32 \times 32$  images. This is caused by problems with scale invariance in such small images. In additional tests for  $32 \times 32$  images, with random translation and rotation, but without random scaling, the accuracy degradation is not observed (Table 5).

Scale inv. func.: $S^1$			Scale inv. func.: $s^2$		
Tanimoto	Eucldn	Manhttn	Tanimoto	Eucldn	Manhttn
0.867	0.867	0.858	0.891	0.883	0.900

Table 5. Results for  $32 \times 32$  images without scale variance, 1-Nearest-Neighbour, 8 test and 1 sample image per person, Yale faces dataset

In very small images the problems with discretization of the method arise. First, the accuracy of the Radon transform deteriorates, affecting both scale and rotation invariance. Moreover, the part of the scale invariant method involving transition to angle domain from Radon spatial domain becomes instable. The accuracy of the  $\gamma$  coordinate in (8) and (9) decreases. Finally, small image size results in loss of the quality of interpolation, when changing the irregular points along the  $\gamma$  axis into a regular mesh. For small images in the same scale, as in Table 5, these inaccuracies become systematic, and thus do not result in a significant loss in accuracy.

The results for all 11 images from the Yale dataset, including the left- and right-side illuminated, are summarized in Table 6. These results are used to compare the method with other face recognition methods, which used the full Yale dataset.

	Scale inv. func.: $S^1$			Scale inv. func.: $s^2$		
	Tanimoto	Eucldn	Manhttn	Tanimoto	Eucldn	Manhttn
$256 \times 256$	0.800	0.800	0.873	0.891	0.885	0.879

Table 6. Results for 3-Nearest Neighbour and leave-one-out scheme, 11 images per person, Yale faces dataset

The objects in the above-described tests have boundaries different for each class but similar within each class. Thus, a proof that the invariant methods did not use the shape of the object as the discrimination criterion is needed. To show that the interior of the image is used for discrimination, tests for the BioID dataset have been carried out. The results are presented in Table 7. It should be noted that the images in this set are not translated, rotated or scaled. On the other hand, the images do not contain the same amount of information as Yale images, since the face in the image is closely cropped. Moreover, the shape of the object, i.e. the head, is different for each individual in the Yale dataset. The results for BioID dataset show that the method gives good results even when the boundaries of all objects in all classes are the same. This proves that the proposed invariants discriminate the objects using the information contained within its interior and not by their shapes. It also proves that enough information for recognition is preserved from the image interior throughout the invariant transform.

Scale inv. func.: $S^1$			Scale inv. func.: $s^2$		
Tanimoto	Eucldn	Manhttn	Tanimoto	Eucldn	Manhttn
0.800	0.860	0.800	0.890	0.880	0.885

Table 7. Results for 1-Nearest-Neighbour, 5 sample per person and 200 images test set, BioID faces dataset

## 5.1 Optimal Configuration of the Method

On the basis of the conducted tests, guidelines concerning the optimal variants of the method can be specified. For problems with smooth objects, such as face recognition, the  $s^2$  scale invariant is significantly better than the  $S^1$  invariant. On the other hand, for objects with sharp edges, such as mammogram labels, the opposite is true. This is due to the difference in the function proposed for scale invariance – the derivative used in  $s^2$  is not accurately approximated in digital images with sharp edges.

In the  $s^2$  scale invariant, usually the Tanimoto metric yields better recognition accuracy than the Manhattan and Euclidean metrics. However, the differences are not significant. In case the method is used with a large set of sample images, as in the leave-one-out scheme, the optimal input image size is  $64 \times 64$ . If only a limited set of

sample images is available, larger input images yield better results. In  $S^1$  invariant, the Euclidean metric outperforms both the Tanimoto and Manhattan metrics in cases, where large sets of non-distorted sample images are present. In cases of distortions, e.g. cropped mammogram labels test, or small sample sets, e.g. in some test for faces, the Manhattan metric is superior.

## 5.2 Comparison with Other Methods

The proposed invariant transformation has been compared with several methods for object recognition. The face recognition problem was used as a model problem, as it is one of the most widely studied object recognition tasks.

The method has been compared with two established face recognition methods, Eigenfaces and Fisherface [2]. The nearest-neighbour method with leave-one-out technique has been applied to measure the recognition accuracy on the Yale dataset. In the tests, the proposed method operated on randomly translated, rotated and scaled Yale faces. The results for benchmark methods, which operated on original Yale faces, are cited after [9]. There are two differences in our testing scenario and the one used for the benchmark methods. First, in the benchmark methods the face was pictured on a uniform white background. In the proposed method, the faces have been extracted and placed in a uniform black background. Second, the benchmark methods are not invariant methods. Thus, the faces used for benchmark methods were of constant scale, rotation and position. The purpose of this comparison is to show that the proposed method achieves similar accuracy as existing method, while offering additional benefit of being invariant. Since the Radon-based method does not claim to be illumination invariant, two results for the proposed transformation are presented, for tests with and without the two side-illuminated faces per person, respectively. The results, using the best performing variants of the proposed method, are summarized in Table 8.

Method	Recognition accuracy
Proposed method (no side-illuminated images)	96%
Proposed method	89%
Eigenface [9]	81%
Eigenface w/o 1 <sup>st</sup> three [9]	89%
Fisherface [9]	94%

Table 8. Comparison with other face recognition methods, Yale dataset

## 6 CONCLUSIONS

The paper extends the new Radon-based invariant recognition method, proposed by the author in [1]. The method has been presented in more detail and tested for two object recognition problems. The method is proven to be invariant to image

translation, on-the-plane rotation and scaling. The proposed transformation can be used as a preliminary step providing an invariant feature set for some other non-invariant methods, or as a stand-alone method for object recognition. In the latter case, the method deals successfully with recognition of mammogram labels and faces. In the face domain, the method gives results comparable or even better than some established non-invariant methods, while allowing for recognition of faces pictured in different positions, head rotations and distances from the camera.

### Acknowledgements

The author would like to thank Prof. Dr. Witold Dzwiniel for guidance and Dr. Tadeusz J. Popiela, M.D. for consulting in mammography and radiology. The author is also grateful to Mr. Marcin Kurdziel for his remarks.

### REFERENCES

- [1] ARODŹ, T.: On New Radon-Based Translation, Rotation and Scaling Invariant Transform for Face Recognition. In Marian Bubak, G. Dick van Albada, Peter M.A. Sloot, Jack J. Dongarra, editors, Computational Science – ICCS'2004, Lecture Notes in Computer Science, Vol. 3039, 2004, Springer-Verlag.
- [2] BELHUMEUR, P. N.—HESPANHA, J. P.—KRIEGMAN, D. J.: Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection. *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 19, 1997, No. 7, pp. 711–720.
- [3] BRANDT, R. D.—LIN, F.: Representations That Uniquely Characterize Images Modulo Translation, Rotation and Scaling. *Pattern Recognition Letters* 17, 1996, pp. 1001–1015.
- [4] DAUBECHIES, I.: *Ten Lectures on Wavelets*. 1992, SIAM.
- [5] HU, M. K.: Visual Pattern Recognition by Moment Invariants. *IEEE Trans. Inform. Theory*, Vol. IT-8, 1962, pp. 179–187.
- [6] JESORSKY, O.—KIRCHBERG, K.—FRISCHHOLZ, R.: Robust Face Detection Using the Hausdorff Distance. In J. Bigun and F. Smeraldi, editors, *Audio and Video based Person Authentication – AVBPA 2001*, 2001, pp. 90–95. Springer.
- [7] KHALIL, M. I.—BAYOUMI, M. M.: Affine Invariants for Object Recognition Using the Wavelet Transform. *Pattern Recognition Letters* 23, 2002, pp. 57–72.
- [8] KIM, H.-K.—KIM, J.-D.: Region-Based Shape Descriptor Invariant to Rotation, Scale and Translation. *Signal Processing: Image Communication* 16, 2000, pp. 87–93.
- [9] LAI, J. H.—YUEN, P. C.—FENG, G. C.: Face Recognition Using Holistic Fourier Invariant Features. *Pattern Recognition* 34, 2001, pp. 95–109.
- [10] SHAO, Y.—CELENK, M.: Higher-Order Spectra (HOS) Invariants for Shape Recognition. *Pattern Recognition* 34, 2001, pp. 2097–2113.
- [11] STRANG, G.—NGUYEN, T.: *Wavelets and Filter Banks*. 1996, Wellesley-Cambridge Press.

- [12] SUCKLING, J. et al.: The Mammographic Image Analysis Society Digital Mammogram Database. *Exerpta Medica. International Congress Series* 1069, 1994, pp. 375–378.
- [13] SUJAN, V. A.—MULQUEEN, M. P.: Fingerprint Identification Using Space Invariant Transforms. *Pattern Recognition Letters* 23, 2002, pp. 609–919.
- [14] THEODORIDIS, S.—KOUTROUMBAS, K.: 1999. *Pattern Recognition*. Academic Press.
- [15] DE VEL, O.—AEERHARD, S.: Object Recognition Using Random Image-Lines. *Image and Vision Computing* 18, 2000, pp. 193–198.
- [16] WOOD, J.: Invariant Pattern Recognition: A Review. *Pattern Recognition*, Vol. 29, 1996, No. 1, pp. 1–17.



**Tomasz Arodz** received his M.Sc. in computer science at the AGH University of Science and Technology in Kraków (Poland) in summer 2003. Currently he is working towards his Ph.D. degree at the Institute of Computer Science at the AGH. His research interests include pattern recognition methods and their applications in biology and medicine. He focuses on ensemble classifiers applied to computer-aided cancer diagnosis in radiology images and mass spectra, as well as to prediction of anticancer drug activities. He is also involved in research on performance analysis tools for grid computing.