

On New Radon-Based Translation, Rotation, and Scaling Invariant Transform for Face Recognition

Tomasz Arodz^{1,2}

¹ Institute of Computer Science, AGH, al. Mickiewicza 30, 30-059 Kraków, Poland

² Academic Computer Centre – CYFRONET, Nawojki 11, 30-950 Kraków, Poland
gefarodz@cyf-kr.edu.pl

Abstract. The Radon transform has some interesting properties concerning the scaling, rotation-in-plane and translation of the input image. In the paper, these properties are a basis for deriving a transformation invariant to the aforementioned spatial image variations, a transformation that uses direct translation, angle representation and 1-D Fourier transform. As the face images often differ in pose and scale of the face, such a transformation can ease the recognition task. Experimental results show that the proposed method can achieve 96% and 89% recognition accuracy for, respectively, uniformly and non-uniformly illuminated images.

Keywords: Face recognition; Radon transform; Invariant recognition

1 Introduction

In face recognition, one often encounters the problem that pictures of the same person are taken in different conditions. These include variations in facial expression, illumination as well as spatial location and pose of the face in the picture. The latter ones consist of translation and scaling of the face, rotation on-the-plane and rotation in depth. In this paper, a method for compensating translation, scaling and rotation-on-the-plane will be shown. The goal is to present a transformation which will give identical results if applied to a pair of images that are in the similarity relation with each other. The result of such a transformation form a new, invariant set of features for recognition.

Several approaches to the problem of invariant object recognition have been proposed up to date. These include group of methods based on the Fourier and log-polar or the Fourier-Mellin transform [6,8] or Taylor and Hessian invariants [2]. Methods based on algebraic moments, e.g. the Zernike [11] or Hu [3] moments are also proposed. Another approach, using only object boundary has been proposed, e.g. a method utilizing wavelets [4] or multi-vector eigenvector shape descriptors [5]. Other methods, using e.g. a group of random lines through the object are also used [10]. The Radon transform has been used as part of the invariant transform [7], albeit not in the face recognition. In this paper the Radon transform will be used as a preliminary step for deriving the invariance.

The paper is arranged as follows. In Sect. 2 the Radon transform is studied and subsequently the full invariance is derived. Section 3 discusses implementation choices and the experimental results for Yale [1] dataset. Finally, Sect. 4 concludes the paper.

2 Method Details

The proposed method consists of two main steps. First, the Radon transform of the image is calculated. Then, the result of the transform is processed to achieve full similarity invariance, using direct translation, angle representation and 1-D Fourier transform.

2.1 Radon Transform

The behaviour of the Radon transform for translated, rotated and scaled images is presented in this section. The result of the Radon transform RAD of the image $f : \mathbb{R} \times \mathbb{R} \rightarrow [0, 1]$ is a function $g : \mathbb{R} \times [0, 2\pi) \rightarrow \mathbb{R}_+$ defined as:

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

where:

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

Given the original image f in Cartesian ($f(x, y)$) and polar ($f_{\text{polar}}(r, \phi)$) coordinates, the following images and their Radon transforms can be defined [7]:

- translated image: $RAD(f(x - x_0, y - y_0)) = g(s - x_0 \cos \theta - y_0 \sin \theta, \theta)$,
- rotated image: $RAD(f_{\text{polar}}(r, \phi + \theta_0)) = g(s, (\theta + \theta_0) \bmod 2\pi)$,
- scaled image: $RAD(f(\alpha x, \alpha y)) = \frac{1}{|\alpha|} g(\alpha s, \theta)$.

The translation in any direction in spatial domain leads to translation in the s direction in the Radon domain, with the translation value varying with the θ dimension. The rotation in spatial domain leads to circular translation along the θ axis in the Radon domain. Finally, the scaling along both axes in the spatial domain results in the scaling along the s axis in the Radon domain and scaling of the value of the transform. These properties are depicted in the Fig. 1.

2.2 Achieving Full Similarity-Transform Invariance

The result of the Radon transform is further processed to obtain translation, rotation and scale invariance. To simplify the notation, the $g(s)$ will denote the values of $g(s, \theta)$ for any specified θ in places where it will not lead to confusion.

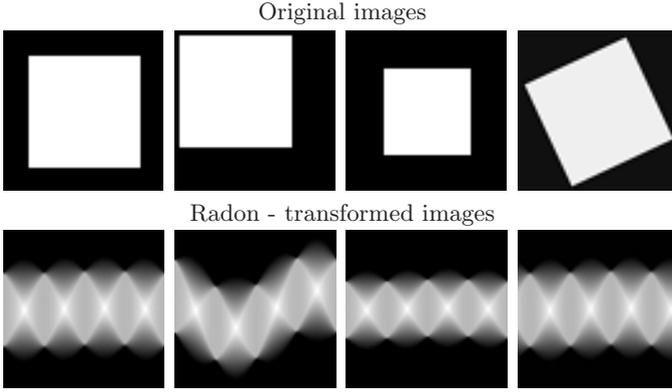


Fig. 1. Effects of the Radon transform for different image transformations

Translation Invariance. To achieve the translation invariance, the result g of the Radon transform is directly translated along the s axis. The value of the translation $s_g(\theta)$ depends on the θ dimension. The resulting function g_t is defined as follows:

$$g_t(s) = g(s + s_g), \quad \text{where} \quad s_g(\theta) = \min \{s : g(s, \theta) > 0\} . \quad (3)$$

The function g_t is translation invariant and preserves the scale variance.

Theorem 1. Let $g(s)$ and $G(\alpha s + S)$ be two functions that are translated and scaled version of each other, i.e. $\alpha g(s) = G(\alpha s + S)$. The functions g_t and G_t , as defined in (3), are scaled version of each other.

The proof of the theorem is straightforward and is omitted for brevity.

Scale Invariance. Scale invariance is based on the method that can be called angle-representation. The process of achieving translation and scale invariance is depicted in Fig. 2.

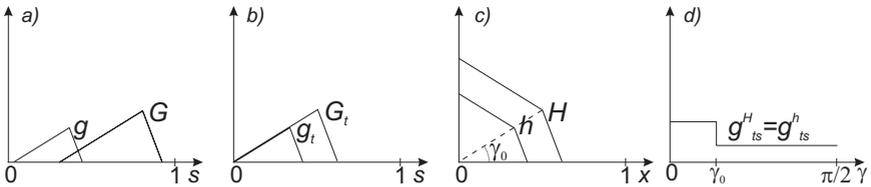


Fig. 2. Achieving invariance to translation: a, b and scale: b, c, d

In the method, a group of functions defined on the basis of the function g_t is used. Let:

$$g_{\text{rev}}(s) = g_t(1-s), \quad \text{where } s \in [0, 1] ; \quad (4)$$

$$h_{\text{rev}}(x) = \int_0^x \left| \frac{dg_{\text{rev}}}{ds} \right| ds, \quad \text{where } x \in [0, 1] ; \quad (5)$$

$$h(x) = h_{\text{rev}}(1-x) . \quad (6)$$

It can be shown that these functions have the following properties.

Theorem 2. *If one assumes that function $g_t(s)$ has continuous first derivative in $[0, 1]$ then the function $h(x)$ is well defined, differentiable, nonincreasing and nonnegative in $[0, 1]$. Also, $h(1) = 0$ and $h(0) > 0$.*

The function h can be used to derive scale invariant transforms. First, let h_P be a variant of the function h narrowed to the domain $(0, X]$, where $X = \min(\{x : h(x) = 0\})$. Let g_{ts}^h be a function $g_{ts}^h : [0, \frac{\pi}{2}) \rightarrow \mathbb{R}_+$ defined as:

$$g_{ts}^h(\gamma) = |h'_P(x)|, \quad \text{where } \gamma = \arctan\left(\frac{h_P(x)}{x}\right) . \quad (7)$$

It can be shown the function g_{ts}^h is scale-invariant.

Theorem 3. *Let $G_t(y)$ and $g_t(x)$ be two functions with the translation removed and meeting the constraints of Theorem 2, such that $G_t(y) = \alpha g_t(x)$ and $y = \alpha x$, i.e two functions that are scaled versions of each other. The function g_{ts}^h is scale invariant, i.e. $g_{ts}^H(\gamma) = g_{ts}^h(\gamma)$, where the functions h and H are defined on the basis of g_t and G_t according to (6).*

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

$$\gamma_H = \arctan\left(\frac{H_P(y)}{y}\right) = \arctan\left(\frac{\alpha h_P(x)}{\alpha x}\right) = \arctan\left(\frac{h_P(x)}{x}\right) = \gamma_h , \quad (8)$$

$$g_{ts}^H(\gamma) = \frac{dH}{d\gamma} = \frac{dH(\alpha x)}{d\alpha x} = \frac{1}{\alpha} \frac{d\alpha h(x)}{dx} = \frac{dh}{dx} = g_{ts}^h(\gamma) . \quad (9)$$

Therefore, g_{ts}^h is indeed scale invariant. □

For clarity, the $g_{ts}^h(\gamma)$, or simply g_{ts} , have been derived for 1-D function $g_t(s) = g_t(s, \theta = \text{const})$, but in fact it is defined in 2-D: $g_{ts}(\gamma, \theta)$ as g_t is.

Rotation Invariance. In order to eliminate the rotation variance, modulus of the one-dimensional discrete Fourier transform is used along the θ axis.

It can be shown [11], that for the discrete function $f : [0, X] \rightarrow R$ the magnitude of the Fourier transform is invariant with respect to circular translation:

$$|DFT(f(x))| = |DFT(f((x + x_0) \bmod X))| . \quad (10)$$

The rotation in the image is reduced to the circular translation in the θ direction by the Radon transform and is preserved in this form by translation and scale variance elimination, both operating along the s axis. Therefore, the function:

$$G_{\text{tsr}}(\gamma, \Theta) = |DFT_{\gamma}(G_{\text{ts}}(\gamma = \text{const}, \theta))| . \quad (11)$$

where G_{ts} is a discrete approximation of g_{ts} , is translation, scale and rotation invariant.

3 Experimental Results and Discussion

The developed invariant transform can be applied to face recognition, either as a pre-processing for methods that are not invariant or as a stand-alone method.

In the latter case, the transform G_{tsr} of input image has to be compared with a set of labelled, transformed sample images. This can be done using the nearest-neighbour decision rule, i.e. the person minimizing the distance is chosen. Several metrics for transformed images has been evaluated, i.e. the Euclidean and Manhattan distance and the Tanimoto dissimilarity measure, defined as the inverse of the Tanimoto similarity measure [9].

It should be noted that apart from spatial variance, other variances, such as illumination, exist in face images. The issue of eliminating the effects of different lighting in images is beyond the scope of this paper, but simple histogram equalization is used as a pre-processing step to compensate for minor lighting variances. Also, simple wavelet-based method is used in some experiments.

The experiments were conducted on the Yale faces dataset [1], consisting of 11 pictures for each of 15 individuals. These images have relatively high resolution and consist of the whole head and neck of a person (see Fig. 3).

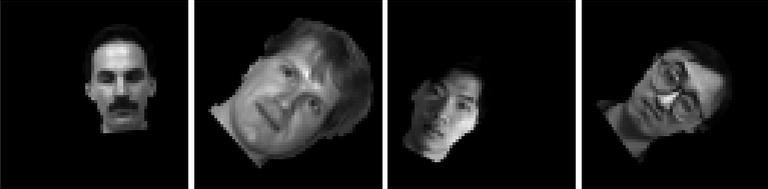


Fig. 3. Faces from the Yale dataset [1]

Table 1. Results for Yale dataset

	3-NN, leave-one-out, 9 imgs/person			8 test and 1 sample imgs/person		
	Tanimoto	Euclidean	Manhattan	Tanimoto	Euclidean	Manhattan
256×256	0.941	0.933	0.919	0.909	0.892	0.892
64×64	0.963	0.956	0.963	0.900	0.867	0.850
32×32	0.919	0.933	0.933	0.775	0.758	0.775

3.1 Experimental Setup

Two pictures for each person from the Yale dataset, containing faces with side illumination has been eliminated from the tests. Two experimental configurations were used. In the first one the "leave-one-out" cross-validation was used to obtain the recognition accuracy. In the second configuration the faces were split into two groups, the sample images database containing 1 image per person and a test set containing remaining 8 images of each person. The picture labelled "normal" in the Yale set, was chosen as a sample image. In both configurations, the nearest-neighbour decision rule was used.

The images consist of a person's head on black background. All heads in the test and sample sets were randomly rotated in the range of $[-\frac{\pi}{2}, \frac{\pi}{2}]$, scaled in the range of [66%, 100%] and placed randomly in the picture.

The images, after histogram equalization, were transformed using the Radon transform to obtain $256\sqrt{2} \times 256$ images. Since the data points in the angle representation after the application of scale-invariance transform are not regularly spaced along the γ axis, piecewise cubic Hermite interpolation was used to obtain regular grid of the size 256×256 pixels. Finally, modulus of the FFT along the θ dimension was calculated.

3.2 Discussion of Results

The results of the tests for the Yale images are summarized in the Table 1. Since the calculation of the Radon transform of the 256×256 image is time consuming, the method has been applied to images downsized to the size 64×64 and 32×32 . While in the nearest-neighbour scenario the reduction of size slightly increases the performance of the recognition, in the scenario with single sample image per person a decrease in accuracy can be observed, especially for 32×32 images. This decrease is caused by the difference in the scale among the images. The relation of the decrease in accuracy to the variations in scale between the faces can be observed with results for 32×32 images randomly rotated and translated, but with no random scaling introduced. For such pictures, the decrease of recognition accuracy is not present, as the recognition accuracy ranges from 0.88 to 0.9 depending on the metric used in the test.

Unlike the translation and rotation, the method for achieving scale invariance utilizes a transformation from the spatial domain to the angle domain. As the size of the image becomes small, the discrete nature of the image becomes more

evident. First, the Radon transform becomes less accurate, which affects also the rotation invariance. Next, the transformation from the Radon spatial domain to the angle domain becomes less stable, e.g. the γ parameter in (7) cannot be determined with good accuracy. Finally, the quality of the interpolation used to change the irregular points along the γ axis into a regular grid deteriorates. While the scale invariance method is used also during the tests with no random scaling, since the scaling in all images is the same, the inaccuracies introduced by the transformation become systematic and do not influence the results significantly. A similar effect can be observed in the nearest-neighbour scenario. Since the number of images that an image is compared with is large, there always exist an image with similar scale of face, for which the errors of transformation are of the same nature as for the tested image. Thus, the recognition rate is not decreased.

Finally, to allow for comparison with other methods, the results for the full Yale set, including two side-illuminated images per person were computed. To deal with the large illumination variance, a wavelet-based method for compensation of illumination effects was used in addition to histogram equalization. This method is labelled HEQ-WV, while the normal histogram equalization is labelled HEQ-64. The results for the images using the nearest-neighbour method are presented in Table 2.

Table 2. Results for 3-NN,leave-one-out, 11 images per person, Yale dataset

	Tanimoto	Euclidean	Manhattan
HEQ-64	0.830	0.824	0.818
HEQ-WV	0.891	0.885	0.879

The results of the tests summarized above allow for some insight into the optimal configuration of the method. Usually the Tanimoto and Euclidean metrics allow for better recognition accuracy than the Manhattan metric. In case of large illumination variance, the wavelet based method for illumination compensation outperforms the histogram equalization. In case the method is used with a large set of sample images, as in the "leave-one-out" method, the optimal input image size is 64×64 . In case of small samples set size, larger input images yield better results.

3.3 Comparison with Other Methods

The proposed method operating on randomly translated, rotated and scaled faces has been compared with results for other methods cited after [6]. Nearest-neighbour method with "leave-one-out" cross-validation technique has been used for the measurement of the recognition accuracy. Two representative methods were used: Eigenface (PCA) and Fisherface. The results are summarized in Table 3. It should be noted that, since the proposed method does not claim to be illumination invariant, the tests of the method were conducted on a test set

with the two images per person, containing right- and left-side illumination, eliminated. Results of the best performing variant of the method, i.e. images of size 64×64 and Tanimoto metric were used. Additionally, for ease of comparison, the results for the whole set, including the two aforementioned images are also presented for the images of size 256×256 and the Tanimoto metric. The results for the benchmark methods are for full Yale dataset, but contrary to the tests of the presented method, the images are not translated, rotated nor scaled.

Table 3. Comparison with other methods using Yale dataset

Method	Recognition accuracy
Proposed method (no side-illuminated images)	96%
Proposed method	89%
Eigenface [6]	81%
Eigenface w/o 1st three components[6]	89%
Fisherface [6]	94%

4 Conclusions

The new Radon-based face recognition method has been proposed. It has the property of invariance with respect to spatial image translation, on-the-plane rotation and scaling. While allowing for recognition of faces pictured in different positions, rotation and scale, the proposed method gives results comparable or even better than some existing, established non-invariant methods. The proposed transformation can be used as a stand-alone method for face recognition or as a preliminary step providing an invariant feature set for some other non-invariant methods.

Acknowledgements. The author would like to thank prof. Witold Dzwiniel for guidance. The author is also grateful to Mr. Marcin Kurdziel for his remarks.

References

1. Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J., 1997. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Mach. Intell.* 19(7) pp. 711-720
2. Brandt, R.D., Lin, F., 1996. Representations that uniquely characterize images modulo translation, rotation and scaling. *Pattern Recognition Letters* 17 pp. 1001-1015
3. Hu, M.K., 1962. Visual pattern recognition by moment invariants. *IEEE Trans. Inform. Theory*, vol. IT-8, pp. 179-187

4. Khalil, M.I., Bayoumi, M.M., 2002. Affine invariants for object recognition using the wavelet transform. *Pattern Recognition Letters* 23 pp. 57-72
5. Kim, H-K., Kim, J-D., 2000. Region-based shape descriptor invariant to rotation, scale and translation. *Signal Processing: Image Communication* 16 pp. 87-93
6. Lai, J.H., Yuen, P.C., Feng, G.C., 2001. Face recognition using holistic Fourier invariant features. *Pattern Recognition* 34 pp. 95-109
7. Shao, Y., Celenk, M., 2001. Higher-order spectra (HOS) invariants for shape recognition. *Pattern Recognition* 34 pp. 2097-2113
8. Sujan, V.A., Mulqueen, M.P., 2002. Fingerprint identification using space invariant transforms. *Pattern Recognition Letters* 23 pp. 609-919
9. Theodoridis, S., Koutroumbas, K., 1999. *Pattern Recognition*. Academic Press,
10. de Vel, O., Aeberhard, S., 2000. Object recognition using random image-lines. *Image and Vision Computing* 18 pp. 193-198
11. Wood, J., 1996, Invariant pattern recognition: A review. *Pattern Recognition*, Vol 29. No. 1, pp. 1-17