

# Complex Object Recognition Based on Multi-shape Invariant Radon Transform

Ghassen Hammouda<sup>1(✉)</sup>, Atef Hammouda<sup>2</sup>, and Dorra Sellami<sup>1</sup>

<sup>1</sup> National Engineering School of Sfax, University of Sfax, Sfax, Tunisia  
gassenhammouda7@hotmail.fr, sellamimasmoudidorra@yahoo.com

<sup>2</sup> The Sciences Institute of Tunis, University of Tunis, Tunis, Tunisia  
atef\_hammouda@yahoo.fr

**Abstract.** Based on the properties of Template Matching and Radon Transform, a new Multi-Shape Invariant Radon Transform (MSIRT) is proposed in this paper. Unlike Radon Transform, integrating projections across lines, the MSIRT uses arbitrary given curves, which are derived from primitives contours. MSIRT leads to peaks once similar shapes are met in the projected image. For seek of genericity and invariance with respect to geometric transformations, we consider different primitives derived from MPEG7 dataset. Each object undergoes a series of pre-processing steps, segmentation and contour extraction, for generating the corresponding primitive. For each query object, the MSIRT is applied with respect to the different primitives and a vote approach will be used for object recognition. Validation of the proposed approach is done on the MPEG7 dataset, giving an accuracy of 94%. Comparison with some known approaches demonstrates the effectiveness of the proposed approach in detecting complex objects, even under geometric transformations.

**Keywords:** MSI Radon Transform · Primitives · Complex objects · Template matching · Vote

## 1 Introduction

Complex-shaped object detection is still an open problem in computer vision. Several works from the literature focus on the detection of objects with common geometric forms (line, square, circle) or parametric forms such as parabolas and hyperbolas. Only few approaches deals with detecting complex objects. However, most of them fail under geometric transformations.

In this paper, we introduce a new formalism for the generalisation of the Radon Transform to detect objects with complex shapes. By building a set of variable primitives, we made our approach invariant to geometric transformations. The remainder of this paper is organized as follows: Related works are described in Sect. 2. The proposed MSI Radon Transform is presented in Sect. 3. Experimental validation on MPEG 7 database and comparison results are given in Sect. 4. Finally, conclusions and perspectives are drawn in Sect. 5.

## 2 Related Works

### 2.1 Template Matching Approaches

Template matching finds out appearance similarities between some template primitives and objects in the image. It ends at potentially locating template shapes in the image. Based on a template illustrating the most relevant traits of appearance of a focused pattern, a matching rate is computed to estimate the occurrence of the considered pattern in a set of images. It is a computational approach that has to deal with possible change in position, scale and rotation or any transformation in the image. The choice of the templates depends on the context and the constraints. To detect the similarity between a template image and a query image with equal dimensions, the cross-correlation approach can be adequate. This approach consists in summing the pairwise multiplications of corresponding pixel values of the images. However, one drawback of the cross correlation is that it cannot handle the change of brightness. Normalized cross-correlation (NCC) [1] is then introduced to improve the original approach. It subtracts the mean image brightness from each pixel value. NCC was adopted to recognize similar forms with a high precision but it is still sensitive to any change of scale or rotation.

### 2.2 Radon Transform

The Radon Transform (RT) is one of the oldest approach. In the literature, several variants of Radon Transforms have been developed [2,3]. Let  $f$  be a function defined on the Euclidean space. Each pixel has a  $(x, y)$  coordinate in a two dimensional cartesian system. So, the Radon Transform can be defined by:

$$R(x', \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x' - x \cos(\theta) - y \sin(\theta)) dx dy \quad (1)$$

Where  $\delta$  is the Kronecker delta function that converts the two-dimensional integral to a line integral along the axis  $x \cos(\theta) + y \sin(\theta) = x'$  and  $\theta$  is the angle of orientation. Radon Transform offers a multitude of properties useful in resolving pattern recognition problems. The most relevant ones in the object recognition are:

- Symmetry:

$$R(x', \theta) = R(-x', \theta \pm \Pi) \quad (2)$$

- Periodicity:

$$R(x', \theta) = R(x', \theta + 2k\Pi) \quad (3)$$

where  $k$  is integer.

- Translation: a translation of  $f$  of  $\bar{w} = (x, y)$  implies a translation of  $\varpi = x_0 \cos(\theta) + y_0 \sin(\theta)$

$$R(x', \theta) = R(x' - x_0 \cos(\theta) - y_0 \sin(\theta), \theta) \quad (4)$$

- Rotation: A rotation of the image by an angle  $\theta_0$  implies a shift of the Radon Transform in  $\theta$ .

$$R(x', \theta) = R(x', \theta + \theta_0) \quad (5)$$

- Scaling: a zoom of  $\alpha \neq 0$  in  $f$  involves a change of scale in Radon Transform:

$$R(x', \theta) = \frac{1}{\alpha} \times R(\alpha * x', \theta) \quad (6)$$

Radon Transform can be useful in pattern recognition. The projection of a pattern with RT is done without loss of information because only the non-null pixels are projected in the Radon matrix in order to retains the relevant information. The RT is also robust against noise. In fact, it can detect some scattered pixels without lack of accuracy. The relevant information detected as straight lines appears as a peak in the Radon space. Indeed, RT performs well in the detection of lines. Rojebani [4] propose an approach for object recognition called the GR-signature (GR). It is essentially based on the Radon Transform and the Gradient to measure the rectangularity of the form. This transform is robust to noise and it is discriminant even under deformation. It allows to estimate the shape of the object based on its characteristics.

S. Tabbone *et al.* [5,6] proposed an hybrid approach called the Histogram of the Transformed Radon (HTR). By statistically analysing the Radon Transform, this approach can detect lines. In other way, it offers a 2D histogram representing the length of the shape given at each direction. The HTR is invariant to translation and rotation but it still very sensitive to any noise or occlusion and detect exclusively lines.

The previous transforms are essentially concerned with straight lines in images. Recently, some works have focused on more complex shapes such as the Polynomial Discrete Radon Transform (PDRT) [7]. The PDRT offers the advantage of projecting a polynomial shape equation in all directions of an image to find it. The sum of pixels of the detected shape will be stored as a peak in the Radon space. In fact, this approach is limited to polynomial curves. The Generalized Radon Transform (GRT) [8] was also defined to project a 2D function over parametrized curves and provides a general solution for some complex forms and it is useful to detect parameterized shapes in an image. However, the GRT suffers from the absence of the multi directional criteria depriving the shape to be detected in different orientations.

To deal with this limitation, Elouedi *et al.* proposed the Generalized Multi-Directional Radon Transform (GMDRT) [9]. It allows to recognize multiple complex geometric curves presented as parametric equation such as circles, rectangles and parabolas in all directions. The GMDRT detects curves with any orientation of the initial shape. Even if the GMDRT offers a significant amelioration in the detection of geometric curves, it remains unable to detect any complex forms since there is no available parametric explicit description for these curves.

The application of various Radon Transform approaches has shown its efficiency in detecting straight lines and geometric forms with rectilinear shape.

An extension of the Radon Transform based on a parametric equation is used to identify the curve of different forms belonging to the same family. It improves the characteristics of the Radon Transform as a shape descriptor. However, application of Radon Transform was often considering specific forms as parabolas, polynomials, etc. The goal of the proposed approach called the Multi-Shape Invariant Radon Transform (MSI Radon Transform) here is to detect complex objects without need of a predefined parametric modeling. This is made possible by applying the MSI Radon Transform of the searched object on a number of primitives to detect its presence.

### 3 The MSI Radon Transform

#### 3.1 General Brief Description

The MSI Radon Transform is a novel approach joining both features: Radon Transform and Template Matching. On the one hand, Radon Transform genericity inherited from considering variable primitives, and on the other hand accuracy of the Template Matching. The result of the application of MSI Radon Transform is some peaks in Radon space, in case of presence of specific shapes in the image. For seek of invariance, we consider in building the primitives different positions and sizes of the objects, we apply geometric transformations (scaling and rotation) to the different images in the dataset. In this way, MSI Radon Transform is made efficient under scaling and rotation. In a next step, similarity between images is computed from Radon space, the obtained peaks are analyzed for affecting each object to its correct class. All these steps are drawn in Fig. 1.

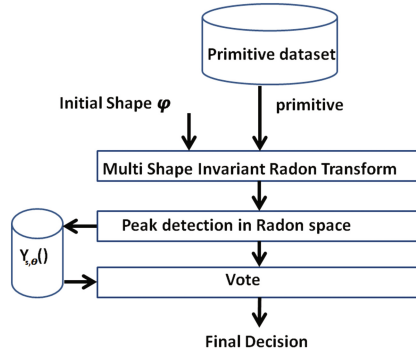
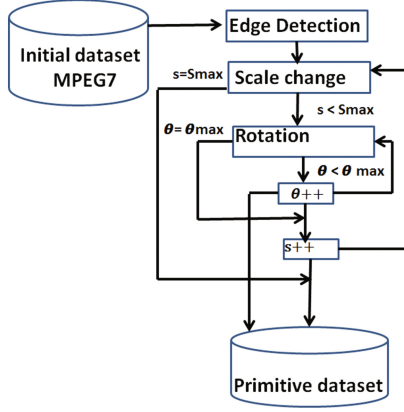


Fig. 1. The MSI Radon Transform based object detection steps

#### 3.2 MSI Radon Transform Formalism

An MPEG7 dataset is used in validation. Let  $\varphi$  be an input initial primitive without any hypothesis made on its shape or size. Each image from the initial

dataset is noted  $I_i$  and represents a two dimensional matrix which have undergone geometric transformations and deformations.



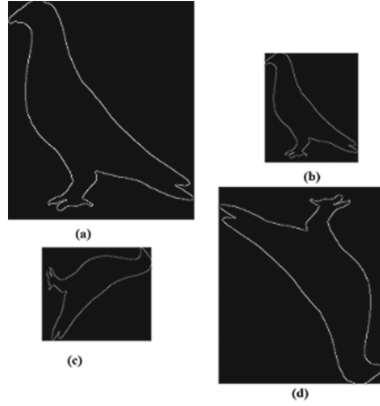
**Fig. 2.** Primitive generation steps

**Primitive Generation.** As shown in Fig. 2, for each Image  $I_i$  in the dataset, we apply a series of preprocessing steps including edge detection, scale change  $s$  and orientation  $\theta$ , for sweeping them.

- Edge Detection: In order to reduce the computation complexity and to focus on the object shape, a contour extraction process is applied. The image is converted into a perceptual space HSV and the Split and Merge technique is applied. This process eliminates the shadow and keeps only the object relevant information. The Canny edge detector operator is then applied. Once the edge is extracted, a binary image is generated.
- Scale Change and Rotation: We apply a scaling of the image by a factor  $s$  ranging from  $s_0 = 0.5$  to  $s_{max} = 2$ . For each scaled image of the  $k$  images of the dataset, we apply rotations by  $\theta$  ranging from  $\theta_0 = 0^\circ$  to  $\theta_{max} = 180^\circ$ . Let  $ns$  be the number of the scaling factors ( $ns = 16$ ) and  $n\theta$  the number of the rotations applied for each scale ( $n\theta = 181$ ). The resultant images  $I_{s,\theta}$  from these iterations constitute a bigger dataset of  $k \times ns \times n\theta$  primitives.

Figure 3 illustrates some primitives generated from a bird image from the dataset MPEG7.  $I_{s,\theta}$  is the result of the preprocessing steps and is given by Eq. (7).

$$I_{s,\theta} = \begin{bmatrix} I_{s,\theta}(-L, 0) & I_{s,\theta}(-L, j) & \dots & I_{s,\theta}(-L, n-1) \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ I_{s,\theta}(0, 0) & I_{s,\theta}(0, j) & \dots & I_{s,\theta}(0, n-1) \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ I_{s,\theta}(L, 0) & I_{s,\theta}(L, j) & \dots & I_{s,\theta}(L, n-1) \end{bmatrix} \quad (7)$$



**Fig. 3.** Some primitives of a bird image from the dataset (a) original primitive (b) processed primitive rotated by  $\theta = 0^\circ$ , scaled by  $s=0.5$  (c) processed primitive rotated by  $\theta = 90^\circ$ , scaled by  $s=0.5$  (d) processed primitive rotated by  $\theta = 180^\circ$ , scaled by  $s=0.9$ .

**MSI Radon Transform.** The MSI Radon Transform is given by:

$$y_\theta(n) = \sum_{m=-M}^{m=M} R_{m,\theta} \times I_{s,\theta}(n+m) \quad (8)$$

$y_\theta(n)$  is the resultant column of the matrix  $y_\theta$  where  $I_{s,\theta}$  the matrix of the primitive corresponding to an angle  $\theta$  and the scale  $s$  starting on the column  $n$  is projected over  $\varphi$ .  $I_{s,\theta}(n+m)$  is a fixed column of  $I_{s,\theta}$ .  $R_{m,\theta}$  are  $(2L+1) \times (2L+1)$  selection matrices introduced by Beylkin where are stored elements of  $I_{s,\theta}(n+m)$  involved in the projection  $y_\theta(n)$  [12]. Each row  $j$ ,  $-L < j < L$  in  $R_{m,\theta}$  store the pixels from  $I_{s,\theta}(n, m)$  belonging to  $\varphi$  starting at the position  $(j, n)$ . The construction of the  $R_{m,\theta}$  consists in presenting the shape of  $\varphi$  in a  $k$  position with  $-L < k < L$ .  $y_\theta(n)$  is then the column resulting in the projection of  $\varphi$  starting in an initial coordinate  $(j, n)$  over the matrix of the primitive  $I_{s,\theta}$ . Each component  $y_\theta(j, n)$  of this column is the sum of the pixels centered on the shape and started in the coordinate  $(j, n)$ .  $M$  represents the number of columns  $I_{s,\theta}$  involved in the computation of  $y_\theta(n)$ .

**Peak Detection.** Values of Radon peaks  $y_\theta(n)$  are stored. They are arranged in a decreasing order for the further vote step.

**Vote.** Once the highest peaks are collected for each primitive, a vote is then used. The object class is then taken as the major class in the first primitives.

## 4 Experimental Results

In this section, an experimental set-up is provided in order to evaluate the performances of the MSI Radon Transform in complex form object detection. A comparison is done with the MPEG7 dataset.

### 4.1 MSI Radon Transform Performance Evaluation

To evaluate the MSI Radon Transform, a sequence of steps are undertaken and interact as a complex pattern recognition process. Below a brief description of the dataset is presented.

**MPEG7 Datasets.** The MPEG-7 standard Core Experiment CE-Shape-1 Part B [10, 11]: Similarity based Retrieval dataset is available for the research community and is composed of 1400 images. In this dataset, 70 classes of different shapes are included with 20 images for each class. These images contain objects with complex forms. Figure 4 illustrates some images in this dataset.



Fig. 4. MPEG7 dataset.

**Metric of Evaluation.** The detection accuracy is used as a metric of evaluation. It is computed with the following equation:

$$R = \frac{TP}{TP + FN} \quad (9)$$

where TP is the total of the relevant images retrieved associated correctly to its original class and FN is the total of the objects affected to the wrong class. Each object of the dataset is compared to all the other objects of the other classes. The TPR also called sensitivity is the ratio of the true detected objects belonging to a specific class. The area under the curve is also used as a metric of evaluation. It is a common evaluation metric for binary classification problems used in order to evaluate a classifier. The area under the curve will be close to 1 in the case of a good classifier.

**Comparison Result.** To evaluate the performance of the MSI Radon Transform in the recognition of complex shape object, this approach is tested in the MPEG7 objects dataset. Moreover, comparison with other existing approaches is also achieved here in order to situate the proposed approach. For each approach, the recognition rate of a set of object forms in the dataset is estimated. All the previous approaches are implemented and represent each object by a shape descriptor specific to the approach used and is classified accordingly. The obtained accuracy rate is used as a metric of evaluation. Comparison Results as illustration, the sensitivity of some classes using several approaches is summarized in Table 1.

**Table 1.** Sensitivity and accuracy for some objects from the MPEG7 database.

	Apple	Circle	Fork	Symbol	Hammer	Fish	Glass	Bottle	Device	Results
GR [4]	0.15	0.1	0.1	0.4	0.2	0.95	0.2	0.45	0.35	0.33
RT [2]	0.9	0.95	0.55	0.95	0.65	0.7	0.7	1	0.8	0.78
GMDRT [9]	0.9	1	0.15	0.25	0.8	0.85	0.75	0.85	0.7	0.72
NCC [1]	0.95	1	0.8	0.9	0.95	1	0.75	1	0.75	0.91
MSIRT	0.95	0.95	0.9	0.9	0.95	0.9	0.95	0.9	0.85	0.94

The analysis of Table 1 reveals that the MSI Radon Transform approach and the NCC present the best detection rates (94% and 91% respectively) with a slight advantage for the MSI Radon Transform. These results concern all the forms evaluated and confirm the consistency of the proposed approach to distinguish any object with an acceptable recognition rate. The analysis by family of primitives for the nine classes described in Table 1 shows a stability of the results for each object class for this approach. For the other approaches (GR, RT, GMDRT), the results vary considerably from one primitive to another. This is the case of GMDRT which gives an acceptable rate for some primitives but is very limited in the recognition in other ones. Moreover, the MSI Radon Transform approach has a great ability to recognize irregular shapes. This is the case of the object fork where most approaches have provided a very low rate in its recognition while the MSI Radon Transform recognizes it at a rate of 90%. Although the results provided by the NCC are fairly close to MSI Radon Transform, it faces problems of scaling and orientation change. Indeed, the results are significantly affected by rotation and scale variations.

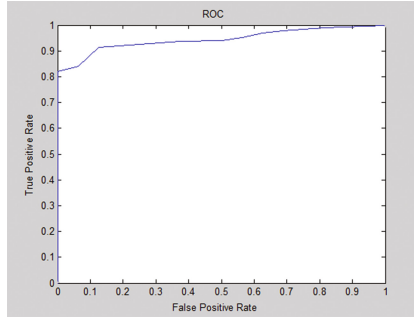
Comparative results given in Table 2 confirm that MSI Radon Transform approach remains stable against rotation and scale variation. The NCC performances are very low illustrating the sensitivity of this approach with respect to changes in these two parameters. However, MSI Radon Transform has some limitations that can be inherited from the contour detection approach and in case of bad contour detector, performance results can be tremendously affected. This latter is affected by the change of rotation and scale and can constitute a kind of limitation. This is the case of the object bottle for example that has



**Table 2.** Comparative study illustrating the performance of the proposed MSI Radon Transform and NCC in detecting primitives of the dataset with scale and rotation change.

	Apple	Circle	Fork	Symbol	Hammer	Fish	Glass	Bottle	Device	Results
NCC [1]	0.85	0	0.8	0	0.1	0	0	0.05	0	0.18
MSIRT	0.95	0.95	0.9	0.9	0.95	0.9	0.95	0.9	0.85	0.94

the lowest rate of 85% caused by the lost of some information in the contour detection. To illustrate the performance of the classifier in the MSI Radon Transform approach, the Receive Operating Characteristic (ROC) is used. We get the curve after sweeping the threshold separating between inter-class and intra-class distributions.



**Fig. 5.** Receive Operating Characteristic curve of the proposed approach.

Figure 5 illustrate an area under the curve (AUC) of 0.94. It denotes the ability of the approach to separate between objects.

## 5 Conclusion

A novel approach for the detection of objects has been proposed. It is a kind of Radon Transform. This transform focuses on the detection of complex shapes objects under geometric transformations changes. A dataset of primitives is obtained by applying preprocessing steps of edge detection, scale and orientation changes on the initial images (here the MPEG7 dataset). The MSI Radon Transform is applied for each query image in order to detect the presence of an initial input object in the dataset. A matrix of peaks in the Radon space revealing a possible presence of a primitive is set and a final vote allows to decide about the right object class. Experiments have been carried out. An area under the curve of 0.94 is obtained. Comparison results show also that the proposed approach outperformed existing ones, by presenting more accuracy and robustness to geometric transformations.

## References

1. Raghavender Rao, Y., Prathapani, N., Nagabhooshanam, E.: Application of normalized cross correlation to image registration. *Int. J. Res. Eng. Technol.* **05**(3), 12–16 (2014)
2. Hasegawa, M., Tabbone, S.: Amplitude-only log radon transform for geometric invariant shape descriptor. *Pattern Recogn.* **47**(2), 643–658 (2014). doi:[10.1016/j.patcog.2013.07.024](https://doi.org/10.1016/j.patcog.2013.07.024). Elsevier
3. Tabbone, S., Wendling, L., Salmon, J.-P.: A new shape descriptor defined on the radon transform. *Comput. Vis. Image Understand.* **102**(1), 42–51 (2006). doi:[10.1016/j.cviu.2005.06.005](https://doi.org/10.1016/j.cviu.2005.06.005). Elsevier
4. Rojebani, H., Elouedi, I., Hamouda, A.:  $R\theta$ -signature: a new signature based on radon transform and its application in buildings extraction. In: *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, pp. 490–495. IEEE (2011)
5. Hasegawa, M., Tabbone, S.: Histogram of radon transform with angle correlation matrix for distortion invariant shape descriptor. *Neurocomputing* **173**, 24–35 (2016). doi:[10.1016/j.neucom.2015.04.100](https://doi.org/10.1016/j.neucom.2015.04.100). Elsevier
6. Hasegawa, M., Tabbone, S.: A shape descriptor combining logarithmic scale histogram of radon transform and phase-only correlation function. In: *International Conference on Document Analysis and Recognition (ICDAR)*, pp. 182–186. IEEE (2011)
7. Ines, E., Dhikra, H., Regis, F., Amine, N.-A., Atef, H.: Fingerprint recognition using polynomial discrete radon transform. In: *4th International Conference on Image Processing Theory, Tools and Applications (IPTA)*, pp. 1–6. IEEE (2014)
8. Hendriks, C.L.L., Van Ginkel, M., Verbeek, P.W., Van Vliet, L.J.: The generalized radon transform: sampling, accuracy and memory considerations. *Pattern Recogn.* **38**(12), 2494–2505 (2005). doi:[10.1016/j.patcog.2005.04.018](https://doi.org/10.1016/j.patcog.2005.04.018). Elsevier
9. Elouedi, I., Fournier, R., Nait-Ali, A., Hamouda, A.: Generalized multidirectional discrete radon transform. *Sign. Process.* **93**(1), 345–355 (2013). doi:[10.1016/j.sigpro.2012.07.031](https://doi.org/10.1016/j.sigpro.2012.07.031). Elsevier
10. de Oliveira, A.B., da Silva, P.R., Barone, D.A.C.: A novel 2d shape signature method based on complex network spectrum. *Pattern Recogn. Lett.* **63**, 43–49 (2015). doi:[10.1016/j.patrec.2015.05.018](https://doi.org/10.1016/j.patrec.2015.05.018). Elsevier
11. Latecki, L.J., Lakamper, R., Eckhardt, T.: Shape descriptors for non-rigid shapes with a single closed contour. In: *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 424–429. IEEE (2000)
12. Beylkin, G.: Discrete Radon transform. *IEEE Trans. Acoustics Speech Sign. Process.* **35**(2), 162–172 (1987)