

# ROTATION, SCALE AND TRANSLATION-INVARIANT SEGMENTATION-FREE SHAPE RECOGNITION

*Hae Yong Kim and Sidnei Alves de Araújo*

Escola Politécnica, Universidade de São Paulo, Brazil  
{hae,saraujo}@lps.usp.br

## ABSTRACT

In this paper, we propose a RST-invariant (Rotation, Scale, Translation) segmentation-free shape recognition method. Most of the existing shape recognition techniques require segmentation before extracting shape features and recognizing it. Unfortunately, segmentation is usually prone to error and segmentation errors produce recognition errors. The proposed technique is based on circular and radial sampling spaces, two 3D spaces built by projecting the analyzed image on circles and radial lines. First, we demonstrate the robustness of the technique in one-scale RT-invariant shape recognition for noisy binary images. We also show that the technique can categorize similar shapes into classes. Then, we make the technique to become invariant to scale. Finally, we demonstrate how the technique can recognize shapes in noisy grayscale images with inconstant background. We demonstrate that, under certain assumptions, the technique is 100% accurate.

## 1. INTRODUCTION

This paper considers RST-invariant (Rotation, Scale, Translation), segmentation-free shape recognition in both binary and grayscale images. This problem occurs naturally in computer vision applications: the vision algorithm must search a noisy image with inconstant background for a query shape that can be darker or lighter than the background, and that can be in any location, any angle and within some range of scales. The “brute force” solution of this problem would be to perform a series of correlations (or template matchings) between the analyzed image and the query shape rotated by every possible angle, scaled by every possible factor (within the scale range) and translated to every possible position. Clearly, this takes too long to be practical. We propose a technique to substantially accelerate this searching, without compromising the accuracy.

To escape from the brute force algorithm, a typical shape recognition algorithm first separates the shape from the background, then extracts some RST-invariant features and compares them with the features of the sample shapes. In the literature, there are many papers on RST-invariant shape descriptors. One of the most important is a set of moments introduced in 1962 by Hu [1]. In recent years, many other techniques that use invariant moments have

been developed [2, 3]. There are other approaches for the shape recognition, for example the curvature scale space proposed by Mokhtarian [4] was adopted by MPEG-7 as standard shape descriptor. Other approaches use circular or radial masks [5, 6]. These techniques are not segmentation-free. Segmentation is usually prone to error, and segmentation error causes recognition error. A segmentation-free RT-invariant system was proposed in [7], but it is not S-invariant and can distinguish only simple shapes. A segmentation-free character recognition technique was proposed in [8], but it is not RS-invariant.

This paper proposes a solution to this problem. It is based on Circular Sampling Space (CiSS) and Radial Sampling Space (RaSS), two 3D spaces built by projecting the analyzed image on circles or radial lines. We show that, under some assumptions, the proposed technique can be as accurate as the brute force algorithm.

## 2. CIRCULAR AND RADIAL SAMPLING SPACES

In this paper, a shape is a binary image defined inside a circle. That is, a query shape  $Q$  is a function  $Q: D \rightarrow \{0,1\}$ , where the domain  $D$  is a circle (figure 1). A shape may be disconnected (figure 1a) or present holes (figure 1c). The aim of this paper is to search an analyzed image  $A$  (binary or grayscale) for a query shape  $Q$ . The shape can appear anywhere inside  $A$  and it can be rotated and possibly also scaled. As the shape recognition will search only for the shape inside the domain circle, the center of the domain and its radius may be modified to specify which subpart of the pattern is to be searched for (figures 1a and 1b).

Given a 2D image  $A: \mathbb{R}^2 \rightarrow \mathbb{R}$  to be analyzed, its circular sampling space (CiSS) is a function  $C_A: \mathbb{R}^2 \times \mathbb{R}^+ \rightarrow \mathbb{R}$  defined:

$$C_A(x, y, r) = \int_0^{2\pi} A(x + r \cos \theta, y + r \sin \theta) d\theta$$

Intuitively,  $C_A(x, y, r)$  is the average grayscale of the pixels of image  $A$  situated at distance  $r$  from pixel  $(x, y)$ . A computer graphics algorithm for drawing circles, as [9], can be used to find efficiently all the pixels that belong to a specific circle.

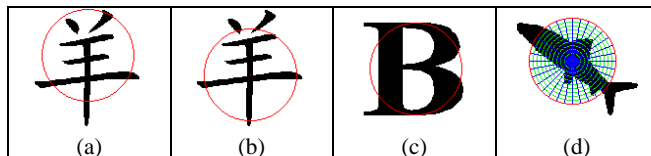


Fig. 1: Examples of shapes. Figure (d) depicts CiSS in green and RaSS in blue.

---

This work was supported by CNPq, process numbers 475155/2004-1 and 307193/2006-3. The authors thank prof. Farzin Mokhtarian for permitting us to use the SQUID image database.

Similarly, given a 2D image  $A$ , its radial sampling space (RaSS) with length  $l$  is a function  $R_A^l: \mathbb{R}^2 \times \mathbb{R}^+ \rightarrow \mathbb{R}$  defined:

$$R_A^l(x, y, \alpha) = \int_0^l A(x + t \cos \alpha, y + t \sin \alpha) dt$$

Intuitively,  $R_A^l(x, y, \alpha)$  is the average grayscale of the pixels of  $A$  located on the radial line with one vertex at pixel  $(x, y)$ , length  $l$  and inclination angle  $\alpha$ . A line drawing algorithm, as [10], can be used to find efficiently all the pixels that belong to a line. Figure 1d illustrates CiSS and RaSS concepts.

### 3. ONE-SCALE CISS BINARY IMAGE ANALYSIS

In the following, we describe the RT-invariant recognition of a shape  $Q$  in a binary image  $A$  at a fixed scale using CiSS. First, CiSS  $C_A(x, y, r)$  is computed for all pixels  $(x, y)$  of  $A$ , and for some predefined set of radii  $r_A = \{r_0, r_1, \dots, r_{K-1}\}$ . In this case,  $C_A(x, y, r)$  is a 3D image where, for each pixel  $(x, y)$ , there is an associated vector with  $K$  features. Then, a central pixel  $(\bar{x}, \bar{y})$  of shape  $Q$  is chosen (usually the center of mass) and the CiSS  $C_Q(\bar{x}, \bar{y}, r)$  is computed only at  $(\bar{x}, \bar{y})$ , for a set of radii  $r_Q$ . In general,  $r_Q \subseteq r_A$ . However, to simplify the explanation, let us assume  $r_Q = r_A$ .  $C_Q(\bar{x}, \bar{y}, r)$  is a feature vector with  $K$  RT-invariant features. We use the mean absolute difference of vectors  $C_Q(\bar{x}, \bar{y}, r)$  and  $C_A(x, y, r)$  to recognize the query shape. An instance of shape  $Q$  is considered to occur at pixel  $A(x, y)$  if the CiSS difference

$$CD_{A,Q}(x, y) = \frac{1}{K} \sum_{k=0}^{K-1} |C_A(x, y, r_k) - C_Q(\bar{x}, \bar{y}, r_k)|$$

is below some threshold  $\varepsilon$ .

Figure 2 illustrates this process. Figure 2a is the analyzed image  $A$ . It contains 18 instances of 7 different shapes rotated by random angles. Some shapes are touching each other and noise and lines were inserted to demonstrate the robustness of the technique. The CiSS of this image was computed for radii  $r = 0, 1, \dots, 47$  pixels. The CiSS of the 7 query shapes (at their central pixels) were also computed. For each one of the 7 query shapes, the pixels where the CiSS difference is below  $\varepsilon=5\%$  were detected. These pixels are painted using 7 different colors in figure 2b (they were dilated to improve visibility). The threshold was chosen to eliminate false negatives. However, the processed image contains many false positive errors. There may exist two different shapes with exactly equal CiSS, causing false positives. However, different instances of a same shape always produce the same CiSS (even using few circles) and in noiseless situation it is always possible to obtain zero false negative rate.

### 4. ONE-SCALE RASS BINARY IMAGE ANALYSIS

We describe here RT-invariant shape recognition in a fixed scale using RaSS. RaSS is not R-invariant *per se*, and needs some more processing than CiSS to achieve RT-invariant recognition. First, we must choose an appropriate length  $l$  (usually the radius of the domain circle  $D$ ) and a set of predefined angles  $\{\alpha_0, \alpha_1, \dots, \alpha_{M-1}\}$ . Given the two images  $A$  and  $Q$  as before, RaSS  $R_A^l(x, y, \alpha)$

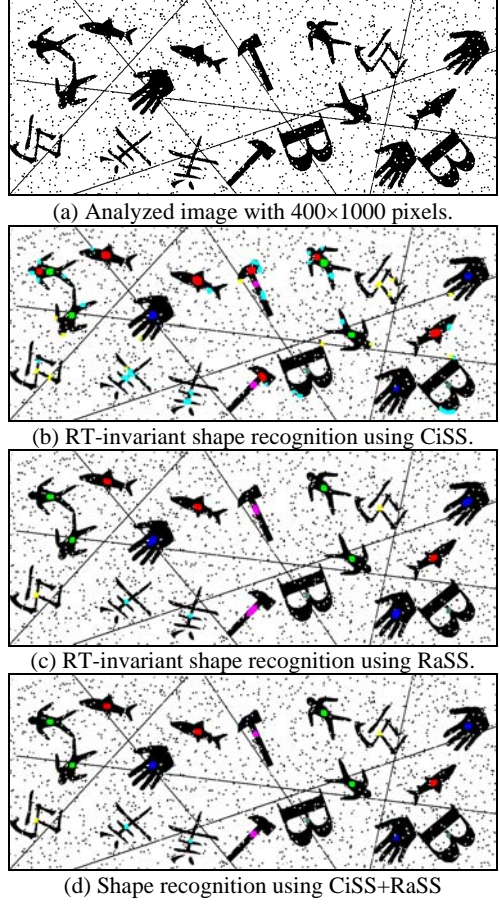


Fig. 2: One-scale shape recognition in binary image.

is computed for all pixels  $(x, y)$  and RaSS  $R_Q^l(\bar{x}, \bar{y}, \alpha)$  is computed only at its central pixel  $(\bar{x}, \bar{y})$ . Then, the RaSS difference at pixel  $(x, y)$  is defined

$$RD_{A,Q}(x, y) = \frac{1}{M} \text{MIN}_{j=0}^{M-1} \left[ \sum_{m=0}^{M-1} |R_A^l(x, y, \alpha_m) - \text{cshift}_j[R_Q^l(\bar{x}, \bar{y}, \alpha_m)]| \right]$$

where “cshift  $j$ ” means circular shifting  $j$  positions of the argument vector. This difference function performs a circular matching, making RaSS features to become RT-invariant.

Figure 2 illustrates this process. The RaSS of  $A$  (figure 2a) and of the seven query images were computed, for 46 uniformly spaced angles and length  $l=47$  pixels. For each one of the 7 query shapes, the pixels where the RaSS difference is below  $\varepsilon=8\%$  were detected. Figure 2c depicts in 7 different colors these pixels (dilated to improve visibility). The recognition contains no errors.

### 5. ONE-SCALE CISS+RASS BINARY IMAGE ANALYSIS

Experimental tests (like the one depicted in figures 2a-2c) show that RaSS provides a better shape discrimination capability than CiSS. However, RaSS shape recognition is more time-consuming than CiSS, because the RaSS features are not naturally invariant to rotation. In order to accelerate the processing, while keeping the RaSS shape discrimination capability, we suggest to compute first the “candidate pixels” where the query shapes can occur using

CiSS (the colored pixels in figure 2b). In noiseless case, the detection of candidate pixels by CiSS can produce false positives, but no false negatives. Then, the RaSS differences  $RD_{A,Q}(x, y)$  are computed only at the candidate pixels, eliminating (or reducing) false positives. The result of this recognition process, named CiSS+RaSS, is depicted in figure 2d. The recognition contains no errors. If images  $A$  and  $Q$  are noiseless, the CiSS+RaSS shape recognition will not produce any false negatives, provided that an enough number of radial lines are used. Indeed, none of our tests contained errors. However, theoretically it is possible that there still remain false positives. In this case, we suggest further filtering the candidate pixels using template matchings. This task is undemanding because RaSS detects the probable shape inclination angle of each candidate pixel. The resulting process can be as accurate as the brute force algorithm at a fixed scale.

## 6. CATEGORIZING SIMILAR SHAPES INTO CLASSES

The CiSS+RaSS technique was also applied to categorize similar shapes into classes (figure 3), instead of detecting exactly equal shapes. The analyzed image contained 62 shapes randomly rotated and randomly chosen from the 7 classes of shapes: fish, quadruped, dude, fighter, hand, ray and tool. Each class contained 3 different shapes. For example, the quadruped class included cat, cow and dog; the fighter class included F16, harrier and skyhawk. The threshold of CiSS was loosened to 7% to allow detecting similar shapes. For each class, a specific RaSS threshold level was empirically defined. The resulting image contained almost no errors: only 3 rays were incorrectly classified as both rays and fighters (two of them are visible in figure 3, painted in yellow and cyan). Similar results were obtained repeating this experience with different seeds of the pseudo-random generator. We cannot expect that the proposed technique outperform more sophisticated (but segmentation-dependent) shape classification techniques. The advantage of the proposed technique is that it does not depend on the segmentation.

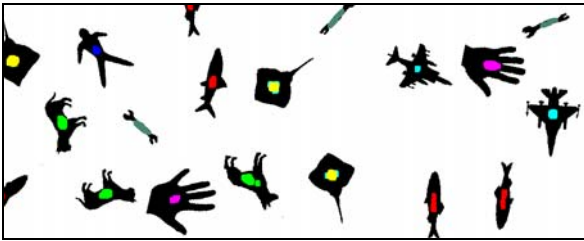


Fig. 3: Categorizing shapes into classes using CiSS+RaSS.

## 7. MULTI-SCALE BINARY IMAGE ANALYSIS

The CiSS+RaSS technique is not intrinsically invariant to scale. However, it is possible to obtain RST-invariant shape recognition with some more processing. The underlying idea is to use the fast CiSS technique to detect the candidate pixels. Associated with each candidate pixel, the probable shape scale is computed. Then, the slow but accurate RaSS technique is used to filter out false positives. If still remain false positives, template matchings can perform an even more accurate shape recognition.

First, it is necessary to delimit a range for shape scales of  $A$ . For example, let us assume that the scales of the shapes in  $A$  can range from 50% to 100% of the size of the query shape  $Q$ . Then,  $Q$  is resampled to various scales inside the specified range, say 50%,

55%, 60%, ..., 100%, yielding the resampled images  $Q_0, Q_1, \dots, Q_{S-1}$ . In these resamplings, the circular domain  $D$  must be resized together with the query shape. The CiSS  $C_{Q_s}(\bar{x}, \bar{y}, r)$  of the scaled query image  $Q_s$  is computed for each scale  $0 \leq s < S$ , yielding  $S$  feature vectors. The number of features  $K_s$  in each vector depends on the scale  $s$ . The CiSS  $C_A(x, y, r)$  of  $A$  is computed as in the one-scale shape recognition. Then, the multi-scale CiSS difference  $MCD_{A,Q}$  is computed for all pixels  $(x, y)$  of  $A$ :

$$MCD_{A,Q}(x, y) = \frac{1}{K} \text{MIN}_{s=0}^{S-1} \left[ \sum_{k=0}^{K_s-1} |C_A(x, y, r_k) - C_{Q_s}(\bar{x}, \bar{y}, r_k)| \right]$$

The probable CiSS scale at pixel  $(x, y)$  is defined:

$$CS_{A,Q}(x, y) = \text{ARGMIN}_{s=0}^{S-1} \left[ \sum_{k=0}^{K_s-1} |C_A(x, y, r_k) - C_{Q_s}(\bar{x}, \bar{y}, r_k)| \right]$$

Pixels  $(x, y)$  where the  $MCD$  is below a threshold level  $\epsilon$  are chosen as the candidate pixels with the respective probable scales.

Figure 4a depicts the result of the searching image  $A$  for the dog shape's candidate pixels (dilated to improve visibility). The shades of color represent the probable scales (red=100% and yellow=50%). Unfortunately, there are many false positives that must be filtered out by the RaSS technique.

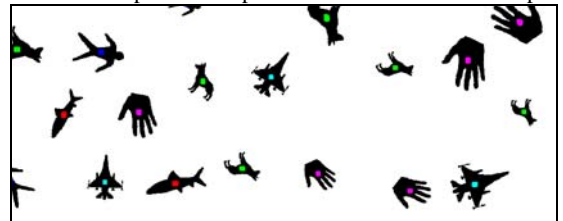
We do not use truly multi-scale RaSS (which would be prohibitively expensive computationally), but apply the RaSS shape recognition only at the probable scale determined by CiSS. To



(a) Dog shape's candidate pixels detected using the multi-scale CiSS recognition. The shades of color represent the probable scales (red=100% and yellow=50%).



(b) Multi-scale CiSS+RaSS recognition of the dog shape. The shades of color represent the probable inclination of the shape.



(c) Multi-scale CiSS+RaSS shape recognition repeated for other shapes. The recognition contains no errors.

Fig 4: RST-invariant shape recognition using CiSS+RaSS.

apply the RaSS technique at a scale  $s$ , the RaSS  $R_Q^l(\bar{x}, \bar{y}, \alpha)$  of  $Q$  is computed exactly as in the one-scale case. The RaSS of  $A$  at scale  $s$  is slightly different from the one-scale definition:

$$R_A^l(x, y, \alpha) = \int_0^{l_s} A(x + t \cos \alpha, y + t \sin \alpha) dt,$$

where  $l_s$  is the length of the radial lines at scale  $s$ . For example, if  $s=70\%$ ,  $l_s=0.7 \times l$ . Then, the RaSS difference at scale  $s$  is defined:

$$RD_{A,Q}^s(x, y) = \frac{1}{M} \text{MIN}_{j=0}^{M-1} \left[ \sum_{m=0}^{M-1} |R_A^l(x, y, \alpha_m) - \text{cshift}_j[R_Q^l(\bar{x}, \bar{y}, \alpha_m)]| \right]$$

The RaSS technique filtered the candidate pixels of figure 4a, yielding figure 4b. The color hues represent the probable shape angles. Figure 4c depicts this technique repeated for other query shapes. The recognition was perfect. As in the one-scale case, this process may produce false positives but it will not produce false negatives, provided that enough numbers of scales and radial lines are used. All our multi-scale CiSS+RaSS tests were error-free. However, theoretically it is possible that there still remain false positives. In this case, it is possible to further filter the candidate pixels using template matchings. This task is undemanding because CiSS and RaSS determine respectively the probable shape scale and inclination. The resulting technique is as accurate as the brute force algorithm.

We obtained the following processing times to search the analyzed image  $A$  with  $800 \times 2000$  pixels for shapes with  $128 \times 128$  pixels, using a 3GHz Pentium-4 computer: 47s to compute CiSS of  $A$  with 24 circles (executed once for all query shapes), 12s to compute a multi-scale CiSS difference (once for each query shape) and 1s to compute the RaSS difference with 46 radial lines at the candidate pixels (once for each query shape).

## 8. GRAYSCALE IMAGE ANALYSIS

In many practical computer vision problems, the analyzed image  $A$  is grayscale: the background color changes throughout the image due to inconstant illumination and diverse environments; the shapes are darker or brighter than the background; and the image is noisy. This situation is emulated in figure 5. Let us assume that the grayscale of pixels of  $A$  ranges from 0 to 1, and  $Q$  is binary (the color is 0 or 1). The CiSS+RaSS shape recognition technique can also be applied to this problem, replacing the mean absolute difference (used to compute CiSS and RaSS differences for binary images) by another difference measure  $D_{aq}$  based on the correlation coefficient. The definition of correlation coefficient between two vectors  $a$  and  $q$  is  $r_{aq} = (\tilde{a}\tilde{q}) / (\|\tilde{a}\| \|\tilde{q}\|)$ , where  $\tilde{a} = a - \bar{a}$  and  $\bar{a}$  is the mean of  $a$ . Correlation coefficient ranges from -1 to +1, and vectors  $a$  and  $q$  must be computed using CiSS or RaSS. A difference measure between the two vectors, independent of the local background color and local contrast, can be evaluated by  $D_{aq} = 1 - |r_{aq}|$ . This equation cannot be applied to regions of  $A$  with almost constant grayscale, because  $\|\tilde{a}\|$  will be almost zero. To avoid divisions by zero, we solve first the linear least squares optimization problem:  $\tilde{a} = \beta \tilde{q}$ , whose solution is  $\hat{\beta} = \tilde{a}\tilde{q} / \tilde{q}^2$ , where  $\hat{\beta}$  is the coefficient that minimizes the error  $(\hat{\beta}\tilde{q} - \tilde{a})^2$ .  $\hat{\beta}$  can be regarded as a measure of contrast between the local background and the shape. Then, we evaluate the difference  $D_{aq}$  only

if the absolute contrast  $|\hat{\beta}|$  is above some threshold. Figure 5 depicts this technique applied to one-scale shape recognition. The result contains no errors.

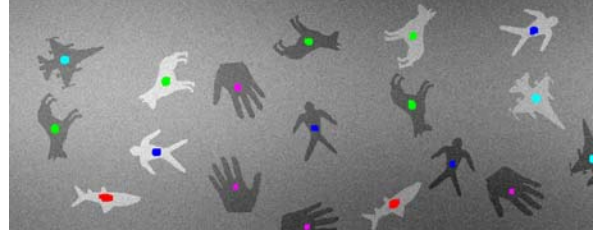


Fig. 5: Grayscale CiSS+RaSS one-scale shape recognition.

## 9. CONCLUSION

In this paper, a segmentation-free RST-invariant shape recognition algorithm was presented. It was applied to analyze both binary and grayscale images. We demonstrated that, under some assumptions, the proposed technique could be as accurate as the brute-force solution. We showed experimentally that the technique is robust to noise and can categorize similar shapes into classes.

## REFERENCES

- [1] M.K. Hu, "Visual Pattern Recognition by Moment Invariants," *IRE Trans. Inform. Theory*, vol. 1, no. 8, pp. 179-187, Feb. 1962.
- [2] J.H. Li, Q. Pan, P.L. Cui, H.C. Zhang and Y.M. Cheng, "Image Recognition Based on Invariant Moment in the Projection Space," *Proc. Int. Conf. Machine Learning and Cybernetics*, Shanghai, vol.6, pp. 3606-3610, Aug. 2004.
- [3] J. Flusser, and T. Suk, "Rotation Moment Invariants for Recognition of Symmetric Objects," *IEEE T. Image Processing*, vol. 15, no. 12, pp. 3784-3790, Dec. 2006.
- [4] F. Mokhtarian, and A.K. Mackworth, "A Theory of Multi-scale, Curvature Based Shape Representation for Planar Curves," *IEEE T. Pattern Analysis Machine Intell.*, vol. 14 no. 8, pp. 789-805, Aug 1992.
- [5] D.H. Chang, and J.P. Hornak, "Fingerprint Recognition Through Circular Sampling," *The Journal of Imaging Science and Technology*, vol. 44, no. 6, pp. 560-564, Dec. 2000.
- [6] Y. Tao, Y. Y. Tang, "The Feature Extraction of Chinese Character Based on Contour Information," in *Proc. Int. Conf. Document Analysis Recognition (ICDAR)*, pp. 637-640, Sep. 1999.
- [7] Y.N. Hsu, H.H. Arsenaault, and G. April, "Rotation-Invariant Digital Pattern Recognition Using Circular Harmonic Expansion," *Applied Optics*, vol. 21, no. 22, pp. 4012-4015, Nov. 1982.
- [8] H.Y. Kim, "Segmentation-Free Printed Character Recognition by Relaxed Nearest Neighbor Learning of Windowed Operator," in *Proc. Brazilian Symp. Comp. Graph. Image Processing (Sibgrapi)*, pp. 195-204, Oct. 1999.
- [9] J.E. Bresenham, "A Linear Algorithm for Incremental Digital Display of Circular Arcs," *Comm. ACM*, vol. 20, no. 2, pp. 100-106, Feb. 1977.
- [10] J.E. Bresenham, "Algorithm for Computer Control of a Digital Plotter," *IBM Systems Journal*, vol. 4, no. 1, pp. 25-30, 1965.