

# GENERIC POLAR HARMONIC TRANSFORMS FOR INVARIANT IMAGE DESCRIPTION

Thai V. Hoang<sup>1,2</sup>, Salvatore Tabbone<sup>1</sup>

<sup>1</sup>LORIA, UMR 7503, Université Nancy 2, 54506 Vandœuvre-lès-Nancy, France  
<sup>2</sup>MICA, UMI 2954, Hanoi University of Science and Technology, Hanoi, Vietnam

## ABSTRACT

A class of rotation-invariant orthogonal moments is proposed using a complex exponential in the radial direction. Each member of this class, while sharing beneficial properties to image representation and recognition like orthogonality and rotation-invariance, has distinctive properties depending on the value of a parameter, making it more suitable for some particular applications. The computation of moments is simpler and more stable than existing methods. Experimental results show the effectiveness of this class of moments in term of description performance and pattern recognition ability.

**Index Terms**— Unit circle orthogonal moments, harmonic radial kernels, rotation invariance, polar harmonic transforms

## 1. INTRODUCTION

Rotation-invariant features of images are usually extracted by using moment methods [1] in which an image  $f(x, y)$  on the unit circle is decomposed into a set of kernels  $V_{nm}(x, y)$  as:

$$\begin{aligned} H_{nm} &= \iint_{x^2+y^2 \leq 1} f(x, y) V_{nm}^*(x, y) dx dy \\ &= \int_0^{2\pi} \int_0^1 f(r, \theta) V_{nm}^*(r, \theta) r dr d\theta, \end{aligned} \quad (1)$$

where  $V_{nm}(r, \theta) = R_n(r)A_m(\theta)$  with  $A_m(\theta) = e^{im\theta}$  [2], an asterisk denoted the complex conjugate, and  $R_n(r)$  could be of any form. For example, rotational moments (RM) [3] and complex moments (CM) [4] are obtained when  $R_n(r)$  has the form  $r^n$ . However, the obtained kernels  $V_{nm}(x, y)$  of RM and CM are not orthogonal and, as a result, information redundancy exists in the extracted moments  $H_{nm}$ , leading to undesirable effects in image reconstruction and recognition. Orthogonality among the kernels means:

$$\begin{aligned} \langle V_{nm}(x, y), V_{n'm'}(x, y) \rangle &= \delta_{nn'} \delta_{mm'} \\ &= \int_0^1 R_n(r) R_{n'}^*(r) r dr \int_0^{2\pi} A_m(\theta) A_{m'}^*(\theta) d\theta. \end{aligned} \quad (2)$$

From the orthogonality of the circular kernels:

$$\int_0^{2\pi} A_m(\theta) A_{m'}^*(\theta) d\theta = \int_0^{2\pi} e^{im\theta} e^{-im'\theta} d\theta = 2\pi \delta_{mm'}, \quad (3)$$

the remaining condition for the radial kernels is:

$$\int_0^1 R_n(r) R_{n'}^*(r) r dr = \frac{1}{2\pi} \delta_{nn'}. \quad (4)$$

There exists a number of methods that satisfies the above condition. One direction is to employ polynomials of the variable  $r$  for  $R_n(r)$ , which turn out to be a special case of the Jacobi polynomials [5]. Popular methods are Zernike moments (ZM) [6], pseudo-Zernike moments (PZM) [3], orthogonal Fourier-Mellin moments (OFMM) [7], *etc* (read [8] for a comprehensive survey). This class of orthogonal moments, despite its popularity, involves computation of factorial terms, resulting in high computational complexity and causes numerical instability, which often limit their practical usefulness.

Another direction is to use harmonic functions for  $R_n(r)$ . The first of this kind is  $R_n(r) = 2 \cos(\pi nr)$ , proposed as an MPEG-7 shape descriptor [9]. The method overcomes the problem of Jacobi polynomials-based moments, but lacks of orthogonality. This problem is finally solved in [10] by using:

$$R_n(r) = \begin{cases} \sqrt{\frac{1}{r}}, & n = 0 \\ \sqrt{\frac{2}{r}} \sin(\pi(n+1)r), & n \text{ odd} \\ \sqrt{\frac{2}{r}} \cos(\pi nr), & n \text{ even} > 0 \end{cases} \quad (5)$$

for radial harmonic Fourier moments (RFHM). The above radial kernels are actually equivalent to  $R_n(r) = \frac{1}{\sqrt{r}} e^{i2\pi nr}$  in terms of image description, similar to the equivalence between different forms of Fourier series. Recently, polar harmonic transforms (PCET) [11] compute image moments by using:

$$R_n(r) = e^{i2\pi nr^2} \quad (6)$$

This paper describes a class of harmonic radial kernels satisfying the condition in Eq. (4) and exploiting the orthogonality of harmonic functions, taking RFHM and PCET as special cases. This leads to a class of transforms, called Generic Polar Harmonic Transforms (GPCET), for the extraction of images' rotation-invariant orthogonal moments. The computation of GPCET is simpler and more stable than existing methods. Each member of this class, while sharing beneficial properties like orthogonality and rotation-invariance, has distinctive property depending on the value of a parameter, making it more suitable for some particular applications. Experimental results show the effectiveness of this class of moments in term of description performance and pattern recognition ability.

The remainder of this paper is organized as follows. Section 2 derives the formula for the generic harmonic radial kernels. Beneficial properties of these radial kernels are presented in Section 3. Experimental results are given in Section 4, and finally conclusions are drawn in Section 5.

## 2. GENERIC POLAR HARMONIC TRANSFORMS

The harmonic radial kernels in Eqs. (5)-(6) are actually special cases of a class of harmonic radial kernels, each defines a set of harmonic radial functions based on which any two different kernels are orthogonal over the unit circle. Indeed, without loss of generality and in order to formulate the generalization, assuming that the harmonic radial kernels has the exponential form  $R_{ns}(r) = \kappa(r)e^{i2\pi nr^s}$ , where  $s \in \mathbb{R}$  and  $\kappa(r)$  is a real functional of  $r$ . Then:

$$\int_0^1 R_{ns}(r)R_{n's}^*(r)r dr = \int_0^1 \kappa^2(r)e^{i2\pi nr^s} e^{-i2\pi n'r^s} r dr. \quad (7)$$

As  $dr^s = sr^{s-2}r dr$ , the right hand side of Eq. (7) becomes:

$$\int_0^1 \kappa^2(r)e^{i2\pi nr^s} e^{-i2\pi n'r^s} \frac{1}{sr^{s-2}} dr^s. \quad (8)$$

By letting  $\frac{\kappa^2(r)}{sr^{s-2}} = \text{const} = C$ , the above formula becomes:

$$\int_0^1 C e^{i2\pi nr^s} e^{-i2\pi n'r^s} dr^s = C \delta_{nn'}. \quad (9)$$

From Eq. (4), it follows that  $C = \frac{1}{2\pi}$  and the generalized definition of the harmonic radial kernels is:

$$R_{ns}(r) = \sqrt{\frac{sr^{s-2}}{2\pi}} e^{i2\pi nr^s}. \quad (10)$$

Consequently, the Generic Polar Harmonic Transforms (GPCET) have the following definition:

$$H_{nms} = \int_0^{2\pi} \int_0^1 f(r, \theta) \sqrt{\frac{sr^{s-2}}{2\pi}} e^{-i2\pi nr^s} e^{-im\theta} r dr d\theta. \quad (11)$$

By taking  $s$  as a parameter, the above definition of  $R_{ns}(r)$  is a true generalization of harmonic radial kernels.  $R_{ns}(r)$  in Eq. (10) becomes RHF in [10] when  $s = 1$  and PCET in [11] when  $s = 2$ . Moreover, by changing the value of  $s$ , a class of harmonic radial kernels is obtained. Members of this class, on the one hand, share beneficial properties to image representation and recognition. On the other hand, each possesses distinctive characteristics determined by the actual value of  $s$  that make it more suitable for some particular applications. These properties will be presented in the next section. Some illustrations of the phase of  $V_{nms}(r, \theta)$  are given in Fig. 1 for  $n = 0, 1, 2$  and  $m = 0, 1, 2$  at  $s = 0.5, 1, 2, 4$  (from top to bottom) using  $(n, m)$  as the horizontal index.

## 3. PROPERTIES

### 3.1. Rotation invariance

Image moments computed by GPCET have an inherent property of rotation invariance. Let  $g(x, y)$  be an image obtained by rotating clockwise  $f(x, y)$  an angle  $\phi$  then  $g(r, \theta) = f(r, \theta - \phi)$ . The GPCET moments of  $g(x, y)$ ,  $H_{nms}^g$ , are related to those of  $f(x, y)$ ,  $H_{nms}^f$ , by:

$$H_{nms}^g = H_{nms}^f e^{-im\phi}, \quad (12)$$

and this is the basis for the derivation of invariants from GPCET moments. The exponential factor  $e^{-im\phi}$  can be eliminated by a magnitude operator. For example,  $|H_{nms}^g| = |H_{nms}^f|$  or even more complex forms as described in [12].

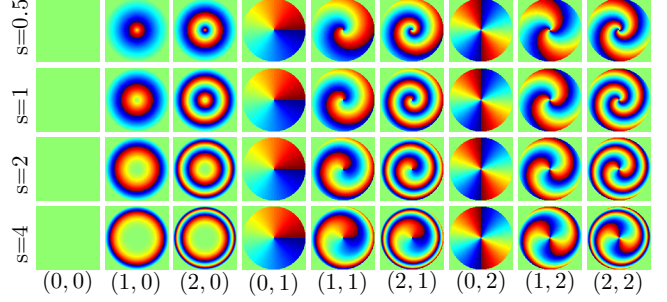


Fig. 1. Some examples of the phase of GPCET kernels.

### 3.2. Zeros of $R_{ns}(r)$

The number of zeros of the radial kernels corresponds to the capacity of moments in describing high frequency components of the images. Since  $R_{ns}(r)$  is a harmonic function of order  $n$ , the equations  $\text{real}(R_{ns}(r)) = 0$  and  $\text{image}(R_{ns}(r)) = 0$  have  $2n$  and  $2n + 1$  distinct roots respectively in the interval  $0 \leq r \leq 1$ . In addition to the quantity, the distribution of these zeros of  $R_{ns}(r)$  is also an important property of radial kernels as it relates to the *information suppression problem* [4]. Suppression is the situation when the computed moments put emphasis on certain parts of the image and neglect the rest. Unlike other methods, the distribution of zeros of  $R_{ns}(r)$  can be controlled by the parameter  $s$  or, in other words, emphasis can be moved to the regions that contain essential discrimination information. When  $s = 1$ , the zeros of  $R_{n1}(r)$  are distributed uniformly, meaning a uniform emphasis over the image region. The more deviation of the value of  $s$  from 1, the more ‘bias’ to the inner (when  $s < 1$ ) or outer (when  $s > 1$ ) parts of the unit circle the distribution of zeros is, corresponding to more emphasis on the inner or outer parts of the images respectively. The suppression can thus be controlled for particular purposes as demonstrated later experimentally. Evidence for these observations on the quantity and distribution of zeros are given in Fig. 2 with the plot of  $\text{image}(R_{ns}(r))$  at  $s = 0.5, 1, 2, 4$ .

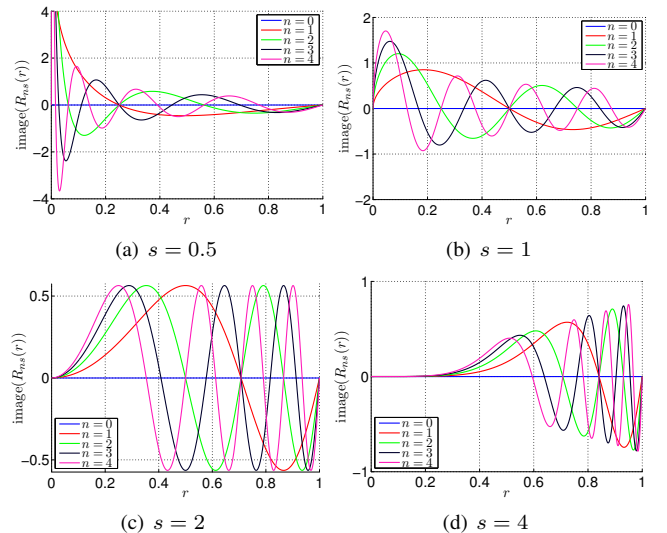


Fig. 2. Imaginary part of some harmonic radial kernels.

### 3.3. Image reconstruction

As the kernels of GPCET form an orthonormal basis over the unit circle, images can be expressed in the following form:

$$f(r, \theta) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} H_{nms} V_{nms}(r, \theta). \quad (13)$$

If only a subset  $\mathcal{S} \subset \{(n, m) \mid n, m = -\infty \dots \infty\}$  of the GPCET moments is available, an approximation of  $f(r, \theta)$  is:

$$\hat{f}_s(r, \theta) = \sum_{(n, m) \in \mathcal{S}} H_{nms} V_{nms}(r, \theta). \quad (14)$$

### 3.4. Fourier series interpretation

GPCET moments can be expressed by means of Fourier series by the following equation:

$$\begin{aligned} H_{nms} &= \int_0^{2\pi} \left[ \int_0^1 f(r, \theta) \sqrt{\frac{sr^{s-2}}{2\pi}} e^{-i2\pi nr^s} r dr \right] e^{-im\theta} d\theta \\ &= \frac{1}{2\pi} \int_0^{2\pi} \left[ \int_0^1 g(r', \theta) e^{-i2\pi nr'} dr' \right] e^{-im\theta} d\theta, \end{aligned} \quad (15)$$

where  $r' = r^s$  and  $g(r', \theta) = \sqrt{\frac{2\pi}{s}} (r')^{\frac{2-s}{2s}} f(\sqrt[s]{r'}, \theta)$ . GPCET moments are thus 2D Fourier coefficients as formulated above: first in the radial direction, followed by the circular direction.

## 4. EXPERIMENTAL RESULTS

Two types of experiments have been carried out to demonstrate the ability of the GPCET moments in terms of description performance and pattern recognition.

### 4.1. Image description

The ability of GPCET in describing image is demonstrated via image reconstruction. In this experiment, a dataset of 26 grayscale images of  $64 \times 64$  pixels of Latin characters in Arial bold font (shown in Fig. 3) has been used. Finite sets of ZM, PZM, OFMM, and GPCET moments are first calculated and then, from these sets of moments, the images are reconstructed using Eq. (14) with  $\mathcal{S} = \{(n, m) \mid n, m < L\}$  for some constant  $L$ . These conditions on the values of  $n$  and  $m$  are selected so that the moments capturing the lowest frequency information are used for the reconstruction process.



Fig. 3. Character images of  $64 \times 64$  pixels in the dataset.

Examples of reconstructed images of GPCET for character E are given in Fig. 4. It is evident that, when  $s = 1$ , the details of E are restored evenly over the image region. When  $s > 1$  ( $s < 1$ ), GPCET moments have difficulty in restoring the inner (outer) part of the images. More deviation results in more difficulty due to the more ‘bias’ in the distribution of the zeros of  $R_{ns}(r)$ . Evidences are reconstructed images of E for  $s = 2$  and  $s = 4$  in the last two rows of Fig. 4.

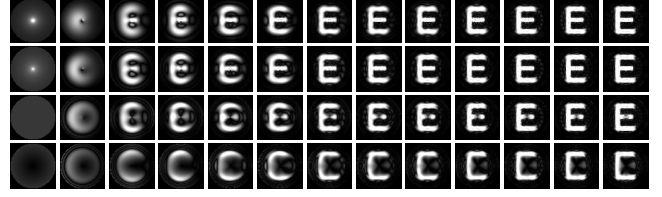


Fig. 4. Examples of reconstructed images of GPCET with  $s = 0.5, 1, 2, 4$  (top to bottom rows) and  $L = 0, \dots, 12$ .

The comparing criteria among different methods is the mean-square reconstruction error (MSRE) [3] defined as:

$$\text{MSRE} = \frac{E \left\{ \iint_{x^2+y^2 \leq 1} [f(x, y) - \hat{f}(x, y)]^2 dx dy \right\}}{E \left\{ \iint_{x^2+y^2 \leq 1} f^2(x, y) dx dy \right\}}, \quad (16)$$

where  $E\{\cdot\}$  is the expectation in ensemble averaging. Fig. 5 shows the MSRE curves of GPCET at different values of  $s$ , from 0.1 to 6 with increment of 0.1, and  $K$  is the number of moments used. It is observed from the color pattern that the value of MSRE decreases slowly when  $s$  has too small or too high value. This is due to the negligence of the extracted moments on certain parts of the images as discussed in Subsection 3.2. Additionally, when  $s < 2$ , there exists a problem of numerical instability due to the existence of the term  $r$  in the denominator of the formula of  $R_{ns}(r)$ . It could thus be concluded that MSRE curves of GPCET initially decrease fast when  $1 \leq s \leq 2$ , have small values in the long run when  $1.5 \leq s \leq 2.5$ , and have no numerical instability when  $s \geq 2$ .

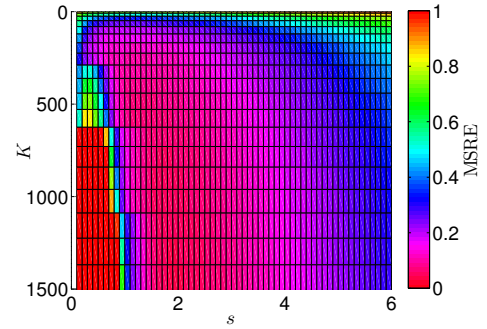
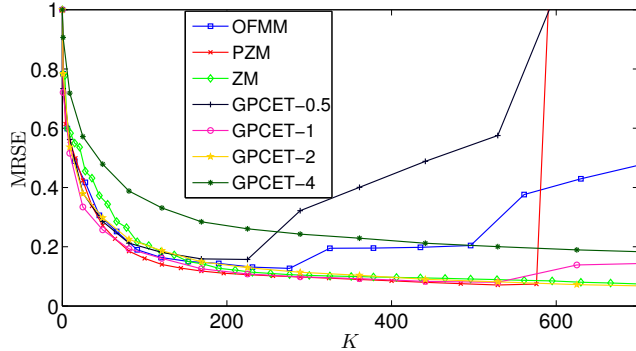


Fig. 5. MSRE curves of GPCET.

Comparison of GPCET with ZM, PZM, and OFMM using MSRE curves is given in Fig. 6. It can be seen that numerical stability of PZM, and OFMM breaks down when the moment order is increased up to a certain value, as evident from the sudden upturn of their respective curve. The curve of ZM breaks down at a higher order (not shown in the figure). The curves of GPCET at  $s = 0.5, 1, 2, 4$  are also plotted and conform with the observations in the previous paragraph.

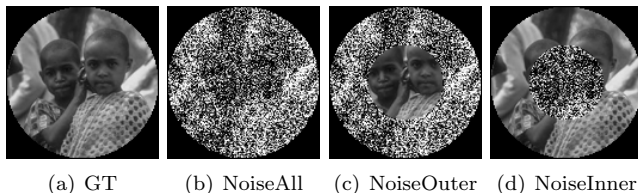
### 4.2. Invariant pattern recognition

This experiment is designed to demonstrate the robustness of GPCET to noise and the role of the parameter  $s$  on recognition results. The experimental images are from the COREL photograph dataset [13]: 100 images have been selected, cropped,



**Fig. 6.** MSRE curves of comparing methods.

and scaled to a standard size of  $128 \times 128$  pixels. These images are the training images, moments extracted from them are taken as the ground truth for comparison with those of the testing images. The testing images are generated from these ground truth images by rotating them with rotating angles  $\phi = 0, 30, \dots, 330$  degrees and contaminating them with Gaussian white noise of variance  $\sigma^2 = 0.2$ . Three different testing datasets NoiseAll, NoiseOuter, and NoiseInner are generated by restricting the noise to be added to the whole image, its outer part, and its inner part respectively (an example is given in Fig. 7). The boundary between the inner and outer parts is the circle of radius 32 having the same center with the image.



**Fig. 7.** An example ground truth image and its three noise-contaminated testing images in three different datasets.

Each noise-contaminated image in these three testing datasets is then classified according to the  $\ell_2$  distance from its moments vector to that of the training images. The average classification rate for ZM ( $K = 10$ ) and GPCET ( $K = 9$ ) at  $s = 0.5, 1, 2, 4$  are given in Table 1. It can be seen that:

- For dataset NoiseAll, the peak performance of GPCET is at  $s = 1$  and, as  $s$  goes away from 1, it decreases.
- For datasets NoiseOuter and NoiseInner, as  $s$  increases from 0.5 to 4, performance of GPCET decreases on NoiseOuter and increases on NoiseInner.
- For all three datasets, the performance of ZM is always lower than the peak performance of GPCET.

The change in performance of GPCET according to  $s$  is predictable using the arguments in Subsection 3.2. GPCET's peak performance on NoiseOuter (NoiseInner) at  $s = 0.5$  ( $s = 4$ ) is due to the bias of the distribution of zeros of  $R_{n,0.5}(r)$  ( $R_{n,4}(r)$ ) towards 0 (1), meaning the information contained in the extracted moments of GPCET is from the inner (outer) part of the images, where the noise is not present. Similarly, peak performance of GPCET on NoiseAll at  $s = 1$  is due to the uniform distribution of zeros of  $R_{n,1}(r)$  over  $0 \leq r \leq 1$ .

**Table 1.** Average classification rate of ZM and GPCET.

	ZM	GPCET			
		$s = 0.5$	$s = 1$	$s = 2$	$s = 4$
NoiseAll	23.75	25.58	30.17	26.00	23.17
NoiseOuter	39.17	58.33	55.58	36.17	28.42
NoiseInner	95.08	79.83	85.25	91.00	98.92

## 5. CONCLUSIONS

This paper has presented a class of rotation-invariant orthogonal moments GPCET using a complex exponential in the radial direction. Each member of this class, while sharing beneficial properties to image representation and recognition, has distinctive distribution of zeros of the radial kernels depending on the value of the parameter  $s$ , making it more suitable for some particular applications. The computation of GPCET is simpler and more stable than existing methods. Experimental results show the effectiveness of this class of moments in term of description performance and pattern recognition ability. Extension of GPCET to 3D data is planned for future work.

## 6. REFERENCES

- [1] M. R. Teague, "Image analysis via the general theory of moments," *J. Opt. Soc. Am.*, vol. 70, no. 8, pp. 920–930, 1980.
- [2] A. B. Bhatia and E. Wolf, "On the circle polynomials of Zernike and related orthogonal sets," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 50, pp. 40–48, 1954.
- [3] C.-H. Teh and R.T. Chin, "On image analysis by the methods of moments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 4, pp. 496–513, 1988.
- [4] Y.S. Abu-Mostafa and D. Psaltis, "Recognitive aspects of moment invariants," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, no. 6, pp. 698–706, 1984.
- [5] Z. Ping, H. Ren, J. Zou, Y. Sheng, and W. Bo, "Generic orthogonal moments: Jacobi-Fourier moments for invariant image description," *Pattern Recognition*, vol. 40, no. 4, pp. 1245–1254, 2007.
- [6] A. Khotanzad and Y.H. Hong, "Invariant image recognition by Zernike moments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 5, pp. 489–497, 1990.
- [7] Y. Sheng and L. Shen, "Orthogonal Fourier-Mellin moments for invariant pattern recognition," *J. Opt. Soc. Am. A*, vol. 11, no. 6, pp. 1748–1757, 1994.
- [8] J. Flusser, T. Suk, and B. Zitová, *Moments and Moment Invariants in Pattern Recognition*, John Wiley & Sons, 2009.
- [9] M. Bober, "MPEG-7 visual shape descriptors," *IEEE Trans. Circuits Syst. Video Techn.*, vol. 11, no. 6, pp. 716–719, 2001.
- [10] H. Ren, Z. Ping, W. Bo, W. Wu, and Y. Sheng, "Multidistortion-invariant image recognition with radial harmonic Fourier moments," *J. Opt. Soc. Am. A*, vol. 20, no. 4, pp. 631–637, 2003.
- [11] P.-T. Yap, X. Jiang, and A.C. Kot, "Two-dimensional polar harmonic transforms for invariant image representation," *IEEE Trans. PAMI*, vol. 32, no. 6, pp. 1259–1270, 2010.
- [12] J. Flusser and T. Suk, "Construction of complete and independent systems of rotation moment invariants," in *CAIP*, N. Petkov and M. A. Westenberg, Eds. 2003, vol. 2756 of *Lecture Notes in Computer Science*, pp. 41–48, Springer.
- [13] J. Z. Wang, J. Li, and G. Wiederhold, "SIMPLcity: semantics-sensitive integrated matching for picture libraries," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 9, pp. 947–963, 2001.