PHASE-BASED METHODS FOR FOURIER SHAPE MATCHING

Paul Fieguth

Systems Design Engineering
University of Waterloo
Waterloo, Ontario, Canada
pfieguth@uwaterloo.ca

Paul Bloore *Adrian Domsa*

Idée Inc.
223 Queen St. E.
Toronto, Ontario
{paul,adrian}@ideeinc.com

ABSTRACT

Interest in the problem of shape matching has been recently reignited by the huge proliferation of images on the Internet, particularly huge computer databases containing thousands or millions of images. Although many typical image database algorithms rely on a variety of shading, texture, and colour attributes, there are significant opportunities for the use of shape as a discriminator, particularly for binary or high-contrast images.

Fourier shape descriptors have been studied extensively for shape comparison, however the descriptor phases have been mostly neglected — either ignored entirely or treated simplistically. This paper formalizes the use of phase in shape matching, and derives shape discriminators, applicable to both simple and complex shapes.

1. INTRODUCTION

The image processing literature has had a long interest in the problem of shape matching; indeed, the definitive references date back almost thirty years [6, 10].

Despite such a long history, interest in the problem has been recently reignited[3, 9, 4, 7, 1, 5] by the fantastic proliferation of electronic imaging and images, particularly huge computer databases containing thousands or millions of images. Although many typical image database algorithms, such as finding two matching or cropped photographs, may rely on a variety of shading, texture, and colour attributes, there are significant opportunities for the use of shape as a discriminator. The problem of trademark infringement is a particularly fitting example: most trademarks are high-contrast binary images (foreground on background) from which a shape is readily acquired, and where shading, texture, and colour play only a secondary role, or none at all.

Fourier shape descriptors have been studied extensively, however the descriptor phases have been mostly neglected. The phase has either been ignored entirely, leading to highly ambiguous matches, or treated simplistically – in a manner which is reasonable for complex, natural shapes, but which fails for simply-structured shapes common among trademarks.

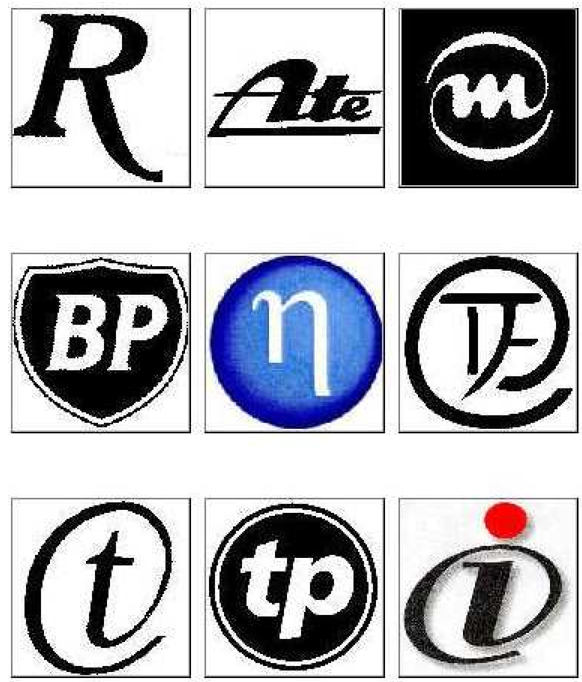

Fig. 1. How do we find matching shapes? There is a very large class of important problems which require the approximate matching of binary images based on their shape or outline.

We emphasize that the purpose of this paper is a theoretical analysis of phase in the context of shape matching. We will derive highly-robust phase tests, invariant to shift, rotation, scale, and orientation. However this leaves us with a *class* of tests, of which the optimum, found through detailed experiments, is the subject of ongoing research.

Section 2 describes shape-matching background, followed by phase models in Sections 3 and 4, and finally a brief discussion.

2. BACKGROUND

Given a huge number of binary images, which are to be tested on the basis of shape, there are two, obvious criteria:

1. Comparison Speed, and
2. Compactness of Representation.

If we restrict ourselves to two-dimensional shapes bounded by simple, closed curves, then it stands to reason that the one-dimensional boundary may be a much more compact description than the huge two-dimensional array of pixels making up the shape.

Indeed, chain codes been developed for such a purpose, in which $x(t), y(t)$ clockwise trace the outline of a shape, where parameter t measures the boundary length from some arbitrary starting point $t = 0$.

Since $x(t)$ and $y(t)$ are clearly periodic, Fourier approaches have long been used[8, 6, 10] in compacting and analyzing shapes. In particular, we let

$$z(t) = x(t) + iy(t) \tag{1}$$

represent the shape outline in the two-dimensional, complex domain; because the images are pixellated, t is not a continuous variable, so we discretize

$$z(n) = x(n) + iy(n) \tag{2}$$

using discrete parameter n (where, in practice, there may be resampling issues, to preserve proper notions of arc-length). Using the FFT, $z(n)$ is easily converted to the Fourier domain

$$f(k) = \mathcal{F}\{z(n)\} \tag{3}$$

where the complex $f(k)$ are known as the Fourier shape descriptors.

The entire challenge, then, of shape matching is the interpretation and comparing of the Fourier descriptors $f_m(k)$ from shape m with those of some other shape n.

Because of the sensitivity of the phase (the complex angle component) of $f(k)$ to image rotations and changes in the shape origin (where we define our "start" point for tracing the outline), often only the magnitude components are examined. However the pitfalls of ignoring phase are long-established in image processing; indeed, each column in Figure 2 has constant magnitude characteristics. The situation becomes progressively worse for more complex shapes, which would include many trademark images, where wildly differing[1] shapes may be naively matched by magnitude comparators.

The next section develops shape phase, followed by approaches to incorporate phase into shape tests.

[1] Anecdotally — this research was started by trying to understand why a simple trademark, consisting of a letter, was best matched (ignoring phase!!) by images of the Andromeda galaxy!

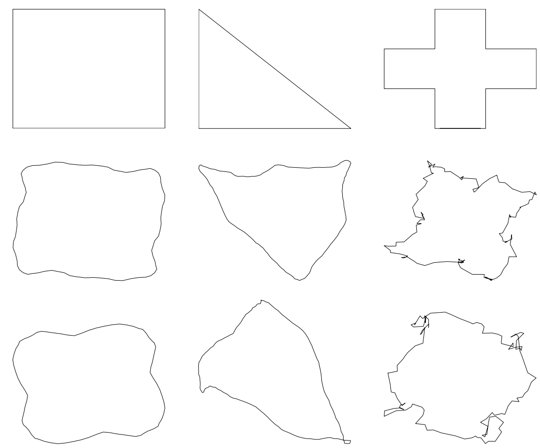

Fig. 2. The sensitivity of shape to the phase information in the Fourier descriptors is obvious from the above three examples: in each column the magnitude information is constant, but with varying phase. Generally, the more complex the shape, the greater the phase sensitivity.

3. PHASE MODEL

Given some model shape outline $z(n)$, we can consider four perturbations which do not affect the inherent shape:

Rescaling: $z_a(n) = rz(n)$
Rotation: $z_b(n) = e^{i\phi} z_a(n)$
Shifting: $z_c(n) = \Delta + z_b(n)$
Reorienting: $z_d(n) = z_c((n-\delta) \bmod N) = \bar{z}(n)$

where N is the length of the boundary, in pixels; thus

$$\bar{z}(n) = \Delta + e^{i\phi} rz((n-\delta) \bmod N). \tag{4}$$

Applying the Fourier transform, we find that

$$\bar{f}(k) = e^{-i2\pi k\delta/N} \left\{ \Delta \cdot (k=0) + e^{i\phi} rf(k) \right\} \tag{5}$$

Since $z(n), \bar{z}(n)$ are equivalent shapes, the goal is to determine how to find descriptors invariant to r, ϕ, Δ, and δ, and thus variant only to inherent changes in shape.

First, the term $\bar{f}(0)$ is the only one sensitive to Δ, so we ignore it. Next, we normalize to $\bar{f}(1)$,

$$\bar{\bar{f}}(k) = \frac{\bar{f}(k)}{\bar{f}(1)} = e^{-i2\pi(k-1)\delta/N} \frac{f(k)}{f(1)} \tag{6}$$

which eliminates the effects of r and ϕ, leaving us with

$$|\bar{\bar{f}}(k)| = \frac{f(k)}{f(1)} \tag{7}$$

$$\angle\bar{\bar{f}}(k) = -2\pi(k-1)\delta/N + \angle f(k) - \angle f(1) \tag{8}$$

The magnitude $|\bar{\bar{f}}(k)|$ is, as desired, independent of the perturbations r, ϕ, Δ, δ, leading to the usual temptation to ignore the phase terms. A more detailed look at phase, motivated by Figure 2, follows.

4. PHASE DISCRIMINATION

It is easy to see that

$$\angle\bar{\bar{f}}(k) - \angle\bar{\bar{f}}(k-1) = -2\pi\delta/N + \angle f(k) - \angle f(k-1) \quad (9)$$

So, ignoring the constant term $-2\pi\delta/N$, the variation about this "mean" should reflect the inherent shape. There are two issues, however, in inferring this mean:

1. Since δ is an angle, and thus periodic, circular statistics must be used in computing a mean; this is straightforward:

$$\hat{\delta} = \frac{-N}{2\pi}\angle\left\{\sum_j \exp\left[i\left(\angle\bar{\bar{f}}(j) - \angle\bar{\bar{f}}(j-1)\right)\right]\right\}. \quad (10)$$

2. If we consider two shapes, $z_1(n), z_2(n)$, which are identical except for the phase in the qth Fourier coefficient,

$$f_1(q) = a\angle\phi_1,\ f_2(q) = a\angle\phi_2, \quad (11)$$

then the mean-square difference between the shapes in the spatial domain, by Parseval equivalent to the difference in the Fourier domain, is

$$\sum_n |z_1(n) - z_2(n)|^2 = \sum_k |f_1(k) - f_2(k)|^2 \quad (12)$$

$$= a^2|e^{i\phi_1} - e^{i\phi_2}|^2 \quad (13)$$

$$= 2a^2(1 - cos(\phi_1 - \phi_2)) \quad (14)$$

That is, the significance of a phase difference is, intuitively, in proportion to the square of the coefficient magnitude. Therefore the common homogeneous treatment of phase differences is clearly *inappropriate*! In the limiting case, if the magnitude $|f(k)|$ is zero, then the phase $\angle f(k)$ is *random*, and thus completely meaningless.

To address the magnitude significance, we can weight the circular mean by the magnitude:

$$\hat{\delta} = \frac{-N}{2\pi}\angle\left\{\sum_j \exp\left[i\left(\angle\bar{\bar{f}}(j) - \angle\bar{\bar{f}}(j-1)\right)\right]|\bar{\bar{f}}(j)\bar{\bar{f}}(j-1)|\right\} \quad (15)$$

This last approach is novel, and generally successful, however it clearly fails if

$$|\bar{\bar{f}}(j)| \cdot |\bar{\bar{f}}(j-1)| \simeq 0 \text{ for all } j \quad (16)$$

Although this may seem highly unlikely – indeed it is, for irregular shapes – it is quite possible for simple shapes, such as a square.

Instead, we need to be able to work with phase differences other than those from adjacent Fourier coefficients. In particular, recalling that

$$\angle\bar{\bar{f}}(k) = \angle f(k) - \angle f(1) - 2\pi(k-1)\delta/N, \quad (17)$$

then the weighted sum of angles is

$$\sum_j \alpha_j\angle\bar{\bar{f}}(k_j) = \quad (18)$$

$$\sum_j \alpha_j(\angle f(k_j) - \angle f(1)) - \frac{2\pi\delta}{N}\sum_j \alpha_j(k_j - 1),$$

which becomes independent of δ if $\sum_j \alpha_j(k_j - 1)$ is an integer multiple of N. For example,

$$\angle\bar{\bar{f}}(3) + \angle\bar{\bar{f}}(N-1)$$

$$= \angle f(3) + \angle f(N-1) - \frac{\delta 2\pi}{N}(2 + N - 2) \quad (19)$$

$$= \angle f(3) + \angle f(N-1). \quad (20)$$

This idea can be generalized to normalize out δ for all coefficients. Given a reference index $j \neq 1$, we can create

$$|\bar{\bar{\bar{f}}}(k)| = |\bar{\bar{f}}(k)| \quad (21)$$

$$\angle\bar{\bar{\bar{f}}}(k) = \angle\bar{\bar{f}}(k) - \angle\bar{\bar{f}}(j)\cdot\frac{k-1}{j-1} \quad (22)$$

$$= \angle f(k) - \frac{k-1}{j-1}\angle f(j) + \frac{k-j}{j-1}\angle f(1) \quad (23)$$

where j is selected to be a meaningful reference; that is, $|\bar{\bar{f}}(j)|$ is not small.

Thus we have constructed a δ-invariant phase sequence! Given two shapes $f_1(k), f_2(k)$, the phase comparison thus involves selecting a good common reference j, typically by maximizing $|f_1(j)\cdot f_2(j)|$, computing $\bar{\bar{\bar{f}}}_1, \bar{\bar{\bar{f}}}_2$, and discriminating as

$$\sum_k |\bar{\bar{\bar{f}}}_1(k)\cdot\bar{\bar{\bar{f}}}_2(k)|\left(2 - 2\cos(\angle\bar{\bar{\bar{f}}}_1(k) - \angle\bar{\bar{\bar{f}}}_2(k))\right). \quad (24)$$

In practice, only some subset (typically at low frequency; e.g., the first $n \ll N$ and the last n) of the phase values may be needed for effective discrimination.

5. CONCLUSIONS

The phase-based discrimination has been implemented for a trademark-matching problem, illustrated in Figure 3. The database of 650,000 images can be searched in a few seconds! This speed is highly significant, since it implies that a user can interact with a search tool in near real-time, adjusting parameters and search criteria on the fly. The performance has been very satisfying, particularly in the elimination of highly irrelevant matches which appeared in earlier algorithms which ignored phase.

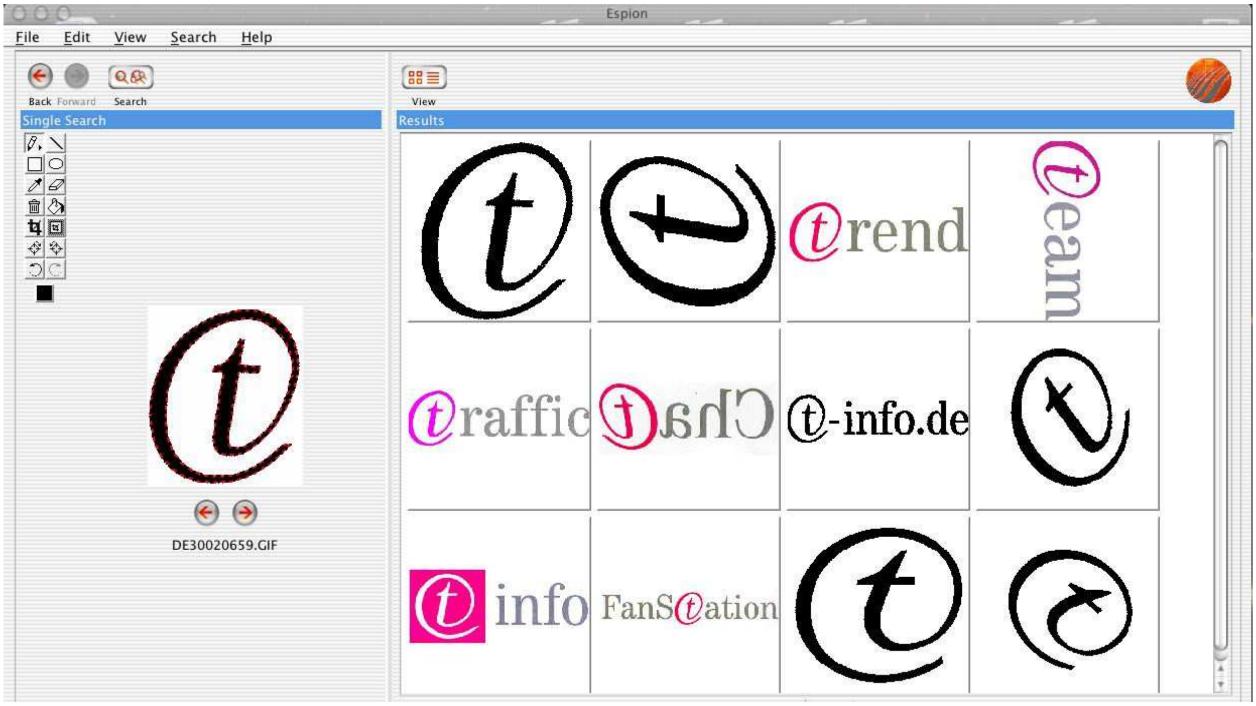

Fig. 3. A sample screen shot from the current search engine prototype, the **Espion** search tool of *Idée Inc.* A database of 650,000 images was searched in a few seconds to produce the above results. As is clear, although the database is huge, only highly-relevant matches are found.

The goal of this paper was not to propose a single shape discriminator, since the details of shape comparison (eg, relative weights of magnitudes and phases, number of phase terms n to include etc.) can be context dependent (and, indeed, as we have seen, may be left to the user for interactive control). Rather, we have demonstrated the clear value in adding the phase context to shape discrimination, and we have derived a robust class of phase-based shape tests, which avoids the pitfalls of random or irrelevant phase in structured shapes.

6. REFERENCES

[1] T. Bui, G. Chen, "Invariant Fourier-wavelet descriptor for pattern recognition," *Pattern Recognition*, 1999

[2] H. Kauppinen, T. Seppanen, M. Pietikainen, "An Experimental Comparison of Autoregressive and Fourier-Based Descriptors in 2D Shape Classification," *IEEE Trans. Pattern Anal. Machine Intell.* (17), pp 201-207, 1995

[3] R. Leon, L. Sucar, "Human Silhouette Recognition with Fourier Descriptors," *International Conference on Pattern Recognition* (3), pp 3713-3716, 2000

[4] W. Liu, T. Wang, H. Zhang, "A Hierarchical Characterization Scheme for Image Retrieval," *Proc. of IEEE Image Processing* (3), pp 42-45, Sep, 2000

[5] W. Ming-Fang, S. Hsin-Teng, "Representation of 3D Surfaces by two-variable Fourier descriptors," *IEEE Trans. Pattern Anal. Machine Intell.* (20), 1998

[6] E. Persoon, K. Fu, "Shape Discrimination Using Fourier Descriptors," *IEEE Trans. Sys. Man, Cyb.* (7), pp.170–179, 1977

[7] B. Pinkowski, "Fourier descriptors for characterizing object contour," *International Conference on DSP Applications and Technology*, pp 1007-1011, Boston, 1996

[8] Y. Rui, A. She, T. Huang, "A modified Fourier Descriptor for shape matching – A Practical Approach," *First International Workshop on Image Databases and Multi Media Search*, 1996."

[9] Y. Rui, A. She, T. Huang, "A Modified Fourier Descriptor for Shape Matching in MARS", *Image Databases and Multimedia Search* (8), World Scientific Publishing, Singapore, pp 165-180, 1998

[10] C. Zahn, R. Roskics, "Fourier Descriptors for Plane Closed Curves," *IEEE Trans. Computers* (21) #3, 1972