Multiscale Image Registration Using Scale Trace Correlation

Bruce B. Hansen
micwatabe@hotmail.com

Bryan S. Morse
morse@byu.edu

Follow this and additional works at: https://scholarsarchive.byu.edu/facpub

Part of the Computer Sciences Commons

Original Publication Citation

BYU ScholarsArchive Citation
Hansen, Bruce B. and Morse, Bryan S., "Multiscale Image Registration Using Scale Trace Correlation" (1999). All Faculty Publications. 615.
https://scholarsarchive.byu.edu/facpub/615

This Peer-Reviewed Article is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Faculty Publications by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen amatangelo@byu.edu.
Multiscale Image Registration Using Scale Trace Correlation

Bruce B. Hansen and Bryan S. Morse
Department of Computer Science, Brigham Young University
3361 TMCB, Provo, UT 84602

Abstract
This paper presents a method for registering images at different magnifications (scales) by treating the problem not only as one of scaling the image coordinates but also as one inherently involving multiresolution information. While some existing methods for multiresolution registration do consider the way the resolution (scale) affects the image, they often consider the image one scale at a time, using geometric properties within that scale. Others use multiscale information, usually to produce more robust results, but only to register same-magnification images (e.g., stereo). A scale trace is the set of values that a single pixel takes on as magnification decreases and the effective point spread function correspondingly increases. As such, scale traces capture information across multiple scales. This paper presents a method that uses correlation of scale traces through multiresolution images to find correspondences between images of the same scene at differing magnifications. Because the method relies solely on the multiscale properties of the images themselves and not on spatial correspondence, it is less sensitive to discrete sampling of potential scale factors. Sample results demonstrate that the method is able to accurately identify both the relative position and relative magnification (scale) of the two images.

1 Introduction

Applications involving multiple images of the same scene at varying magnifications require that these images be registered accurately.

Typical methods where scale is a parameter of the registration process involve mapping one image to another through a transformation involving this scale (and possibly other parameters). It is possible to solve for the relative scale change and other transformation parameters precisely through iterative error-minimization methods, but such solutions require an accurate approximation as a starting point (e.g., [7]). Other methods may not require an accurate starting point but require considerable overlap in the images (e.g., phase-correlation methods such as [6]).

If we seek to find a solution for the relative scale change (and possibly other transformation parameters as well), we also need to consider how accurately we can compare images at different scales. Because magnification causes spatial position to change along with resolution, simple area-based correlation, sum of squared intensity differences, or other spatial comparisons are susceptible to the misregistration caused by misaligned pixels. More importantly, though, these methods may also be susceptible to the inherent difference in blurring between the images.

This paper presents a method that relies on the multiscale (multiresolution) properties of the image itself to find approximate relative scale changes. This method involves scale trace correlation: one-dimensional correlation of the values of a single pixel across multiple scales. Such correlation directly compares different-scale information instead of comparing spatial information at various scales—an important distinction. Inaccuracies caused by error in the estimation of the relative scale changes simply become small translations of smooth one-dimensional signatures rather than errors involving spatial misalignment. Such small translations little affect the correlation.

In order to focus on the multiresolution aspects of registering images at significantly different scales, we limit additional free parameters in the method presented here to include only global translation. The ideas presented here should, however, be extensible to include rotation as well as nonuniform translation and scaling.

2 Scale Spaces and Scale Traces

A scale space [10] is the set of all possible images of a single scene at various levels of blurring (resolution). Analyzing information across multiple levels of resolution can often be used to produce more stable and more general results than can be produced by analyzing a scene at one resolution alone ([10], [8], [3] and many others).

If the multiresolution images are "acquired" (generated) from a base image using a scaled, weighted measurement aperture applied uniformly over the image, a scale space $L(x, \sigma)$ may be written as the convolution of the base image with scaled versions of a single measurement aperture: $L(x, \sigma) = L(x, 0) * G(x, \sigma)$ where $L(x, 0)$ denotes the base image, $*$ denotes convolution, and $G(x, \sigma)$ denotes a measurement aperture with size $\sigma$. 
Figure 1: Scale trace for the upper-leftmost target point in Figure 4b. The trace shows how this point is bright at small scales and darker at larger scales, indicating a bright point near to a darker surround. Finer sampling of scales does not change the fundamental shape of the scale trace, only its sampling accuracy.

In order to avoid artifacts from spurious resolution (temporary increases in sharpness as overall resolution decreases), the unique selection of aperture weights is the Gaussian [1]: 
\[ G(x, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-x^2/2\sigma^2} \]
where the blurring parameter \( \sigma \) is the “scale” of the image.

To maintain uniform information change between successive resolutions, one must sample scale exponentially: 
\[ \sigma_t = \sigma_0 f^t \]
where \( f \) is the scale-change factor from one image to another and \( \sigma_0 \) is the base (highest resolution) scale. Intuitively, this agrees with our normal notions of scale traces of the one-dimensional scale trace—does not otherwise change its fundamental nature. In this way, scale traces are extremely robust as point-signatures across multiple scales.

The number of resolutions at which we sample the scale space also affects scale traces in a straightforward way: the number of discrete scales \( \sigma_t \) affects only the sampling of the one-dimensional scale trace—it does not otherwise change its fundamental nature. In this way, scale traces are extremely robust as point-signatures across multiple scales.

Since we are concerned with relative changes in scale between two images, finding matching points in the images means comparing the scale trace of a point in one image to the shifted scale trace of another (supposedly higher-resolution) image. Let \( t_x^k(\tau) \) denote the scale trace at position \( x \) in image \( k \). Comparing position \( x \) in image 1 to position \( y \) in image 2 with supposed relative scales \( \sigma_\tau \) and \( \sigma_{\tau+\Delta\tau} \) gives comparison measure \( c_{x,y}(\Delta\tau) \). When comparing \( t_x^k(\sigma_\tau) \) to shifted \( t_y^2(\sigma_{\tau+\Delta\tau}) \), correlation seems a natural means for such comparison:

\[ c_{x,y}(\Delta\tau) = \sum_{\tau=0}^{N} t_x^1(\sigma_\tau) t_y^2(\sigma_{\tau+\Delta\tau}) \]
(1)

where \( N \) is the number of scales across which to correlate.
Figure 2: Scale trace correlation example. This image shows the result of correlating the scale trace for the upper-leftmost target point indicated in Figure 4b (see Figure 1) with the scale trace of each point in the lower-resolution image in Figure 4a and selecting the scale shift corresponding to the relative magnification between the two images. The scale trace correlation finds well points with similar multiscale local properties (e.g., other telephone keys).

In other words, the $\Delta \tau$th value in the correlation between two scale traces $t_{1x}$ and $t_{2y}$ indicates the strength of the match between $t_{1x}(\tau)$ and $t_{2y}(\tau+\Delta \tau)$ across all $\tau$. Other comparisons (normalized correlation, sum of squared differences, etc.) could be used, but the effect is the same: comparison of shifted scale traces.

4.1 Using Scale Traces to Match Points

If we fix a single point $x$ in image 1 and use Eq. 1 to calculate the scale-trace correlation between it and every point $y$ in image 2, we get a set of values $c_{x,y}(\Delta \tau)$: a measure of how well each point $y$ in the second image corresponds to point $x$ in the first image with respect to a change in scale corresponding to $\Delta \tau$ magnification steps (where each step has a magnification factor $f$). An example of such comparison is shown in Figure 2. Notice how this target point on one telephone key (upper-leftmost target point in Figure 4b) causes a high response at the relative scale offset for points with similar multiscale local properties (i.e., other keys on the phone).

4.2 Using Higher-Order Signatures

The presentation of scale trace correlation so far in this section has been concerned with only zeroth-order information: the intensity values themselves. The effect of scale-trace correlation using only intensity values is to highlight points that change their intensities in similar ways across scale—typically matching edges well to other edges. (But notice that the scale of the edge doesn’t matter—it finds the match and the scale.) Extending this to first-order scale traces (derivatives in $x$ and $y$ across multiple scales, or gradient magnitude likewise) finds good second-order matches, second-order scale traces find good third-order matches, and so forth.

4.3 Using Multiple Points

While scale traces capture significant local information, multiple points may be required to unambiguously match two images. The spatial relationship between these points should scale with the relative change in scale between the images (Figure 3). So, for scale change $\Delta \tau$, we would compare points $x_1$ and $x_2$ to points $y$ and $y + \sigma_{\Delta \tau}(x_2 - x_1)$. We can combine multiple scale trace correlations by multiplying the strengths of the effective matches:

$$c_{x_1,y}(\Delta \tau) c_{x_2,y + \sigma_{\Delta \tau}(x_2 - x_1)}(\Delta \tau)$$

Suppose that we want to combine a set of $n$ points at positions $x + \Delta x_1, x + \Delta x_2, \ldots x + \Delta x_n$ and compare them to their counterparts in an image with relative scale change $\Delta \tau$: $y + \sigma_{\Delta \tau}\Delta x_1, y + \sigma_{\Delta \tau}\Delta x_2, \ldots y + \sigma_{\Delta \tau}\Delta x_n$. This produces an overall match $C_{x,y}(\Delta \tau)$ between $x$ and $y$:

$$C_{x,y}(\Delta \tau) = \prod_{k=1}^{n} c_{x + \Delta x_k, y + \sigma_{\Delta \tau}\Delta x_k}(\Delta \tau)$$

If we again fix $x$ in image 1 and calculate this match for all $y$ in image 2, we get $C_{x,y}(\Delta \tau) = C_{x,y}(\Delta \tau)$, a measure of the match at scale offset $\Delta \tau$ between $x$ with associated offset points $x + \Delta x_k$ and each point $y$ with appropriately-scaled offset points $y + \sigma_{\Delta \tau}\Delta x_k$.

We can now determine the best overall correspondence between the two images by finding the maximum value in $C_{x,y}(\Delta \tau)$. Specifically, this maximum value indicates the best possible match for point $x$ in a higher-resolution image for both position $y$ and relative scale change $\sigma_{\Delta \tau}$ in a lower-resolution image.

4.4 Example

To illustrate the behavior of the scale-trace correlation algorithm, Figure 4 shows the results of applying the algorithm to a pair of synthetically-generated different-scale images for which the correct result is known.
Figure 4: Results of scale-trace correlation for an image (a, with correct correspondence indicated) and automatically selected targets in a higher-resolution subimage (b). The correct scale $\Delta \tau = 7$ (c) and location (d) are clearly indicated by their values in the correspondence space.

Scale spaces were generated for two images using a scaling factor of $f = 1.1$. 32 scales were generated for the higher-resolution image, and 16 scales were generated for the lower-resolution image. Because the two images differ by a scale factor of 2, the expected match (44 should be found at sampled scale $\tau = \log_{1.1} 2 \approx 7.2725 \approx 7$.

Nine target points (Figure 4b) were automatically selected in the high-resolution image by subdividing the image into nine $3 \times 3$ equal-size regions and within each region finding the point whose scale trace had the highest integrated first derivative—in other words, those points whose scale traces demonstrated the most information content. Each of those points was used in the method described in Section 4.3 to produce a correspondence space $C_\tau(y, \Delta \tau)$. The best match was determined by selecting the maximum value in this correspondence space.

Figure 4c shows the scale trace through the correspondence space for this best match. The peak at $\Delta \tau = 7$ is unambiguous. Figure 4d shows a slice of the correspondence space $C_\tau(y, \Delta \tau)$ for $\Delta \tau = 7$. Nearly all of the space had values near zero except for the small $3 \times 3$ cluster of pixels shown in Figure 4d. The maximum value was found only one pixel from the correct location—easily close enough for more precise iterative methods to use as a starting point.

5 Refining the Registration

In some cases, the initially-selected nine points may not be sufficient to produce the correct registration. The confidence of the match can be detected automatically from the strength of the peak in the combined scale-trace correlations relative to the next-best scale match (see Figure 4d). The confidence and accuracy of the registration can then be improved by using additional comparison points.

We have experimented with both random and automatic selection of additional comparison points. If comparison points are selected from the original nine-part subdivision of the image, the points selected are likely to be near those already used. We have found that random selection of additional points seems to produce better results.

The use of additional comparison points can be seen in Figure 5. In this case, the scale-difference estimate is highly ambiguous after the initial registration with nine feature points (5c). With the addition of 18 more points, the result is now unambiguous (5d). The strength of the peak in the combined scale-trace correlation relative to other potential scale-differences was used to automatically recognize when additional features were needed.

6 Hierarchical Scale-Difference Refinement

As seen in Figure 4, the estimate of the scale difference between the images is accurate only to within the discrete sampling of the scale parameter of the scale space. The registration can also be further improved by refining the estimate of this scale difference. Because the initial registration gives a discrete estimate of the scale difference, we have a limited range of scales in which to further search. Thus, we can refine the scale-difference estimate using a coarse-to-fine strategy in the scale dimension itself.
Figure 6: Hierarchical refinement of the scale estimate. Once the initial best estimate is determined, the range of neighboring scales is resampled with finer precision. By using the square root of the previous scale sampling factor on each iteration, only half of the scales on each iteration need to be computed.

Figure 6 shows how this refinement is performed. The space is first sampled with a scale multiplier of \( f = 1.5 \), a fairly coarse initial sampling of the scale space. After an initial estimate of 7.594 \((1.5^3)\) is found, the range between the adjacent sampled scales \([5.063 \ (1.5^4), \ 11.391 \ (1.5^6)]\) is resampled at a sampling rate of \( f = \sqrt{1.5} = 1.22 \). For robustness, the scale space must be generated beyond just the target range for the scale difference. This process is repeated again, each time using a new sampling factor that is the square root of the previous one.

The following observations can be made of this process:

- By using the square root of the previous sampling factor as the new sampling factor, each newly generated scale space shares every other scale with the previously generated space. Thus, only half of the scales in the new space must be generated with each refinement of the estimated scale difference.

- After the initial estimate, the new best estimate of the scale difference always occurs in the first five discrete scales in the new space (those corresponding to the range being expanded). Thus, not all scales in the space must be searched—the others are used for robustness only.

Eventually the limits of the discrete image limit further refinement. Our experience is that the result can usually be refined to within at least two decimal places, and the limits of the refinement can be recognized by monitoring the clarity of the peak in the combined scale-trace correlation.

Figure 7 shows an example of this process for a synthetically-generated case in which the correct scale difference \((2.0)\) is known. Again starting with a scale-sampling factor of \( f = 1.5 \), the best initial estimate is 2.25. Because the scale-difference estimate is high, the estimated corresponding registration area is too small \((113 \times 113)\).

Figure 7: Refinement of the scale estimate. TOP TO BOTTOM: The initial scale-difference estimate is too large on the first iteration. With refinement, the next estimate is closer but too small. After five iterations, the estimate is accurate to within a single pixel. After seven iterations the estimate is accurate to within 0.26% relative error.

After one iteration of refinement, the estimate has been refined to 1.84. This time, the estimated scale difference is too small, and the corresponding registration area
(139 \times 139) is too large. After five iterations, the estimated scale difference has been refined to 1.982, and the corresponding registration area is only one pixel too large (129 \times 129). After two more iterations, the scale difference has been refined to 1.9948 (a relative error of 0.26%), and the corresponding registration area is correct (128 \times 128).

7 Results

Figure 8 shows the results of using scale-trace correlation to register two images of a building facade. The images were taken at two different positions with two different focal lengths, in each case keeping the camera axis orthogonal to the plane of the building’s wall. Even with only nine feature points, the richness of the scale-trace information allows a correct and unambiguous match.

Figure 9 shows similar results for different-scale images of a glass-covered display case. Again, even with only the initially-selected nine comparison points and despite the glare off of the glass, the result is unambiguous in both scale difference (9c) and spatial position (9d).

Figure 10 shows a different pair of images for the same display case. In this case, the result remains ambiguous because of the ambiguity in the image itself. Observe that because the dominant features in the image (the edges of the pieces of paper) are slightly angled from each other, one can slide the high-resolution image up or down a line between the two edges while simultaneously reducing or enlarging the image, respectively. Scale-trace correlation recognizes this ambiguity and encodes it as a streak of high-possibility matches spatially (10d), but for each possible spatial match the scale of the match is clearly defined (10c).

By detailed examination of the text on the pages, one can find the exact match unambiguously, but this can not be done with a handful of feature points. This is a limitation of any feature-based method. However, because scale-trace registration produces a limited set of positions and scales to further search for the correct registration, a spatial correlation approach can then be used to determine the best match. More importantly, we now know the correct scale for each possible position, thus allowing a spatial-correlation approach to perform more accurately.
Figure 11: Registration using a lower-resolution framing image (a). The higher-resolution image (b) is registered to the medium-resolution image (c) by registering the two to the lower-resolution image, determining the area of potential overlap, and automatically selecting comparison points from that region.

A potentially useful application of multiscale image registration is to use lower-resolution "framing" shots to assist in the registration of multiple higher-resolution images. An example of such an application is shown in Figure 11. Once high-resolution (11b) and medium-resolution (11c) images are individually registered to a lower-resolution framing shot, their respective registrations are used to assist in their registration to each other, even though they only partially overlap. The comparison points are automatically selected from the higher-resolution of the two images and only from the region of overlap (11b). The position of the higher-resolution image in the medium-resolution image is accurately determined, even with only 17% overlap (11c).

8 Conclusion

By comparing scale-shifted versions of multiscale point signatures (scale trace correlation), the multiscale information at that point can be leveraged to produce a robust match for images with both translation and scale differences. The confidence of the match can be determined by the strength of the combined scale-trace correlations, and additional points can be added automatically to improve the results. The scale difference can be further refined by employing a coarse-to-fine refinement strategy in the scale dimension itself.

The essential element of this method is that it compares multiscale information across scales at individual feature points directly. Incorrectly estimating the scale difference causes only small shifting of the scale traces and has only a minor effect on the correlation results, instead of causing both spatial misregistration and differences in blurring. The registration algorithm presented here is one way of using this principle, but the essential idea should be adaptable to other registration algorithms.

References


