A Novel Image Transformation For Solving Complex Image Mapping Problems

Igor V. Maslov
St-Petersburg State Polytechnical University
29 Polytechnicheskaya street
Russia195251, St.-Petersburg
ivm3@columbia.edu

Yulia D. Detkova
St-Petersburg State Polytechnical University
29 Polytechnicheskaya street
Russia195251, St.-Petersburg
detkova@avalon.ru

Izidor Gertner
The City College of New York
138th Street at Convent Avenue, NAC 8/206
USA 10031, New York, NY
gertner@cs.ccny.cuny.edu

ABSTRACT

The paper proposes a novel image transformation called Image Local Response (ILR) that can be used for solving complex image mapping problems. The proposed transformation brings together two approaches based on the pixel value distribution and image features. Image local response is defined as the average value of the difference between the transformed and the original copies of the same image whereby the transformation is small, i.e., the components of the corresponding parameter vector have sufficiently small unit values. The response has a few interesting properties useful in image mapping. The validity of the proposed image transformation is shown on sample complex image mapping problems formulated as the multi-objective piece-wise imaging optimization problem.

Keywords

Image mapping, response analysis, imaging optimization, evolutionary algorithm.

1. INTRODUCTION

Many tasks related to digital image processing deal with comparing (i.e., matching or mapping) images of different types and sizes. Examples of such tasks include e.g., image registration, object or target recognition, and pattern matching. These tasks, in turn, play a pivotal role in many important real world applications like remote sensing, security systems, robotics, computer vision, medical imaging, information fusion, and industrial control.

The approaches that can be used for comparing the images can be divided into two main groups.

1. The first group of methods compares the distributions of the pixel values in the images, either explicitly or implicitly. One of the problems associated with this approach is related to the changing light conditions between the images. In this case, the comparison of the pixel values becomes difficult since no matching pixels can be found. Moreover, the comparison of the different types of imagery, e.g., infrared and real visual images obtained from the different types of sensors (as in multi-sensor image fusion) becomes virtually impossible using this approach.

2. Methods in the second group attempt to find a set of salient characteristics, i.e., features that are common for the compared images. Choosing the appropriate features is by no means a trivial task. It becomes even more difficult if the images are simultaneously misaligned and distorted by some sort of complex geometric transformation, e.g., affine or perspective.

The proposed in the paper image transformation uses the combination of the both abovementioned approaches; the transformation is called Image Local Response (ILR). The concept of ILR is somewhat related to image neighborhood and block operations [Seu00a], [Pit00a], [Ima06a], as well as to the node and edge functions proposed in [Muc98a], although it is based on a fundamentally different idea rooted in Green's functions [Bar89a] and response analysis [Ger02].

The paper is organized as follows. Section 2 gives the definition of Image local response and describes its useful properties. Section 3 discusses sample experimental results of object mapping in the case of geometrically distorted images. Section 4 concludes the paper with the summary of the proposed approach.

2. DEFINITION AND PROPERTIES OF IMAGE LOCAL RESPONSE

In digital image processing, solving image mapping (matching) problem means finding an adequate vector $\boldsymbol{V}$ of parameters defining the unknown transformation $\boldsymbol{A}$ between the images. In its most general form, the sought transformation $\boldsymbol{A}$ can be a fairly complex one, although in many cases it can be represented or approximated with some suitable general affine transformation.

The concept of Image Local Response (ILR) is based on a fairly simple and rational idea: since the mapping problem searches for the unknown transformation $\boldsymbol{A}$, it seems logical to explore the response of the image to this particular type of transformation. This task can be accomplished by mapping a transformed image *Img´* onto self (i.e., onto the original image *Img*), with a sufficiently small transformation vector $\boldsymbol{V_u}$. In accordance with this idea, Image local response R_P at a point P is defined as the value of the difference F between the transformed, *Img´* and the original, *Img* copies of the same image. Here, the transformation $\boldsymbol{A_u}$ at the point P is small, i.e., the components of the parameter vector $\boldsymbol{V_u}$ have sufficiently small unit values.

The simplest way of defining the image difference F is to compute a squared difference of the pixel gray values over some area ω_R, in the following way:

$$F = \frac{\sum_{\omega_R} (g(x', y') - g(x, y))^2}{{\omega_R}^2}, \quad (1)$$

where $g(x,y)$ and $g(x',y')$ are the gray values of the image *Img* in the area ω_R before and after the transformation, correspondingly [Bro98a].

Making the area ω_R sufficiently small has two important implications.

1. The general affine transformation fairly accurately approximates other interesting and plausible image transformations (e.g., perspective) that can be found in real world applications [Ros76a]. This means that one can compute ILR once, i.e., for the affine transformation, and then use the computed values in image mapping with some other, even more complex transformations.

2. The difficulty of mapping images with different pixel value distributions can be significantly mitigated when using ILR since the later maps a particular image onto itself (i.e., onto the same pixel value distribution) within a small area.

Here, the area ω_R is called "response area". For convenience and without loss of generality, a square box $r \times r$ can be chosen as the response area, where r is called "response radius". In the case of the general affine transformation, image response has to be computed for the vector $\boldsymbol{V_u}$ defined by nine parameters: the translations DX and DY along the x- and y-axes; the rotation θ in the xy-plane; the non-isotropic scaling factors SX and SY along the x- and y-axes; the shear SHX and SHY along the x- and y-axes; and the reflections RX and RY about the x- and y- axes. The shaded subarea in Figure 1 shows what part of the small response area near the point P will be changing during the unit transformation, in the case of translation, rotation, and scaling.

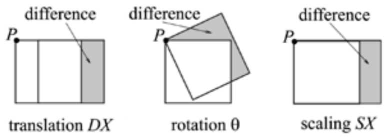

Figure 1. Computing local response at point *P* for translation, rotation, and scaling.

Computing Image local response according to Formula (1) with the chosen small values of ω_R and r is similar to computing Green's functions extensively used in mathematical physics and engineering [Bar89a]. It can be easily shown that, as in the case of Green's function, the response value R_P rapidly decreases as the distance from the point P (i.e., the value of r) increases.

foreach pixel *P*, ***do***
  foreach component of $\boldsymbol{V_u}$, ***do***
    compute **(1)**
  endforeach component
  compute response $R_P = \frac{\sum_{i=1}^{N} F_i}{N}$
endforeach pixel

Figure 2. Algorithm for computing Image local response.

The foregoing definition of the ILR suggests the algorithm shown in Figure 2. In the algorithm, the value of the difference F is computed for each of the N components of the vector $\boldsymbol{V_u}$. In the case of the general affine transformation, $N = 9$. Finally, the response value R_P at the point P is computed as the averaged sum of all N differences F_i ($i = 1, \ldots, N$). In order to compute the response values for the border pixels, the image can be appropriately padded.

The values of image response can be represented in a graphical form - see Figure 3. As one can see, Image local response has a dual nature. On the one hand, ILR is defined in the form of a matrix computed from the pixel value distribution, as Formula 1 suggests. On the other hand, ILR represents the image feature in the form of the contours of the objects that are present in the image. The duality of ILR allows one to transfer the search for the proper image transformation $\boldsymbol{A}$ in image mapping problem from the actual image space $\boldsymbol{I}$ into the response space $\boldsymbol{R}$. In this case, the difference between two images Img_1 and Img_2 can be evaluated as a squared difference of the response values over the area Ω of the overlap of the both images, in the following way:

$$F = \frac{\sum_{\Omega}(R_2(x',y') - R_1(x,y))^2}{\Omega^2}, \qquad (2)$$

where $R_1(x,y)$ and $R_2(x',y')$ are the response values of the reference image Img_1 and the transformed image Img_2, correspondingly.

Figure 3. Original image of a scene (top) and its response representation (bottom).

The different types of imagery are shown in Figure 4 whereby a wireframe and a principal model of the same object expose different pixel values distributions. That makes the mutual mapping of the images with the direct comparison of their gray values impossible. On the other hand, the response images computed according to (1) and shown in Figure 5 exhibit clear definition of the common contours of the both objects, i.e., their main feature.

Figure 4. Original images of an object: the wireframe (left) and the principal (right) model.

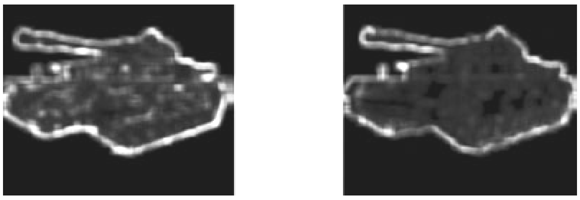

Figure 5. Image local response of the wireframe (left) and the principal (right) model.

Image local response has a few interesting and helpful properties that can be effectively used in computationally intensive image mapping problems.

1. As mentioned before, Image response preserves the main image feature, the contours of the objects in the scene.

2. Using the matrix of the response values significantly reduces the amount of information that has to be processed during the search for the proper transformation $\boldsymbol{A}$. Only the higher response values would participate in the image mapping computations provided the sparse response matrices of the images are represented using efficient data structures.

3. The algorithm for computing ILR shown in Figure 2 can be easily parallelized, so all pixels comprising the image would be processed concurrently on a modern GPU, thus making the computational complexity of the algorithm equal to $O(1)$.

4. Image mapping can be formulated as an optimization problem whereby the image difference plays the role of the objective function that has to be minimized. In this case, ILR provides a smooth bell-shaped fitness landscape very well suited, e.g., for

the evolutionary search where the selection of the successful partial solutions drives the search towards the complete optimal solution [Ash06a].

5. In some cases, the model of Image local response can be effectively used to control local search when image mapping is formulated as an optimization problem. In particular, the value of the vector $\boldsymbol{\alpha}$ = $\{\alpha1, \alpha2, \alpha3, \alpha4\}$ of the coefficients in the Downhill simplex method can be adjusted to the landscape of the objective function thus accelerating the search [Mas05a]. This particular property of ILR is based on the fact that in the close vicinity of the optimal solution, ILR fairly well approximates the objective function, i.e., the global difference between the images.

3. COMPUTATIONAL EXPERIMENTS WITH IMAGE MAPPING AND IMAGE LOCAL RESPONSE

The proposed image transformation in the form of Image local response was tested on a few image mapping problems [Mas08b]. A sample set of three 2D grayscale images is shown in Figure 6. The 300×300-pixel reference image Img_0 contains an object arbitrarily rotated in the 3D coordinate system. Two template images are a 178×195-pixel top view Img_1 and a 185×66-pixel left view Img_2 of the same object. The corresponding image responses computed in accordance with the algorithm given in section 2 are shown in Figure 7.

The search for the proper mapping from the template images onto the reference image is formulated here as an imaging optimization problem solved with a hybrid evolutionary algorithm [Mas08b]. The following conditions are present that complicate the problem:

- two or more template images are used to represent the different views of the same object;
- the object of the mapping undergoes significant distortion caused, e.g., by an arbitrary rotation in the 3D space; such a mapping cannot be defined with a single transformation vector;
- the difference between the images cannot be formulated as a single fitness function; consequently, the search has to deal with the multiple objectives of the optimization.

In accordance with the proposed approach, the search is conducted in the response space $\boldsymbol{R}$, as opposed to the actual image space $\boldsymbol{I}$. In order to accommodate the multiple template images, an advanced computational model is used. The model includes the multiple populations, so that every template is represented by its own independent population. Since the template objects can undergo significant distortion, every template object is divided into k sections, so each section can have its own transformation vector $\boldsymbol{V_k}$. This approach corresponds to a piece-wise approximation of the actual image transformation $\boldsymbol{A}(\boldsymbol{V})$.

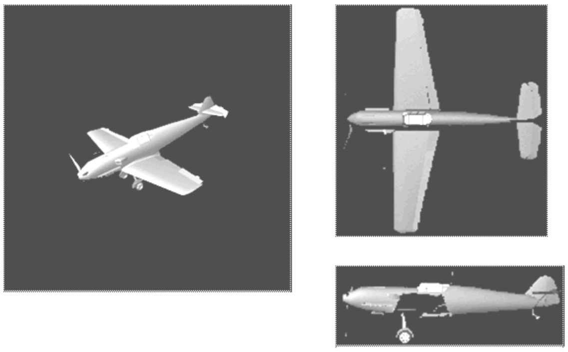

Figure 6. A sample set of three 2D grayscale images: reference image (left) and two template images (right).

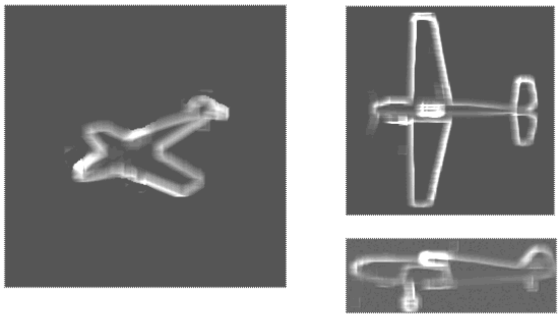

Figure 7. Response images of the sample test set.

The computational algorithm further assumes that every object in the image has some prominent basic feature in the form of a trunk to which all other parts of the object are attached. Here, such a feature is called a “hull”. The transformation of the hull can be defined by the main vector $\boldsymbol{V_A}$ of the general affine transformation and a complementary vector $\boldsymbol{V_D}$ of elastic deformations. The latter vector describes the deviation of the actual hull transformation from the main vector $\boldsymbol{V_A}$.

In its most general form, the entire algorithm works as two relatively independent phases implementing the global search and the local correction. The global search phase attempts to find the optimal solution for the hull transformation, i.e., the best mapping between the template hulls and the reference hull. The local corrections phase attempts to find the optimal piece-wise approximation of the actual image transformation using the hull transformation as its initial approximation. Because of the complex composite structure of the template model and a two-phase search algorithm, one expression for fitness function is not sufficient. The search is conducted in the multi-objective space using the different expressions for the fitness function at the different stages of the algorithm.

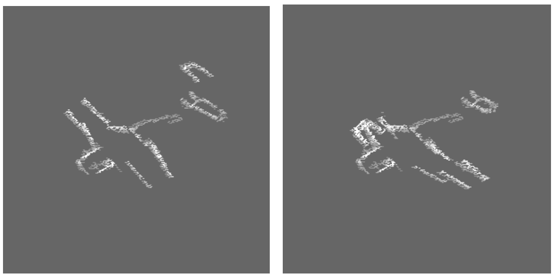

Figure 8. Intermediate results at the different stages of the piece-wise mapping.

Figure 9. Result of the piece-wise mapping of the template objects onto the reference image.

Figure 8 shows some intermediate stages of the piece-wise mapping of the different object sections onto the reference image. Figure 9 shows the final result of the image mapping. As one can see, the template images were successfully mapped onto the reference image using the piece-wise transformations of the original template objects in the response space.

Another interesting and important image mapping problem is medical image registration. Here, different slice images obtained with the CT or MR scan have to be put into the same framework by computing their mutual transformations. Figure 10 presents two sample MR images of different slices. The transformation has to be found that maps the template image Img_1 (Figure 10, right) onto the reference image Img_0 (Figure 10, left).

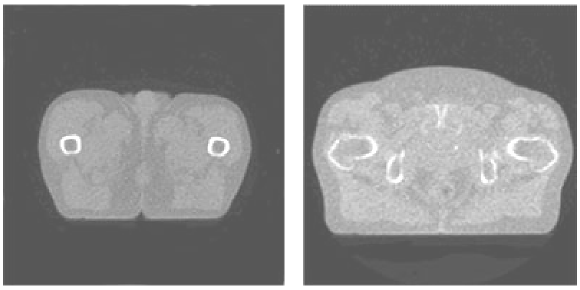

Figure 10. A sample set of MR images: reference (left) and template (right).

The search for the optimal mapping is conducted using the proposed approach, in the same manner as the search for the solution of the previous problem. Figure 11 shows the response matrices of the both images, and Figure 12 presents the final result of the mapping. As one can see, the algorithm was able to find a fairly good mapping of the template onto the reference image. Further improvement of the solution can be achieved with the usage of the adaptive division of images into sections. That would help remove certain roughness and discontinuities in the resulting image transformations.

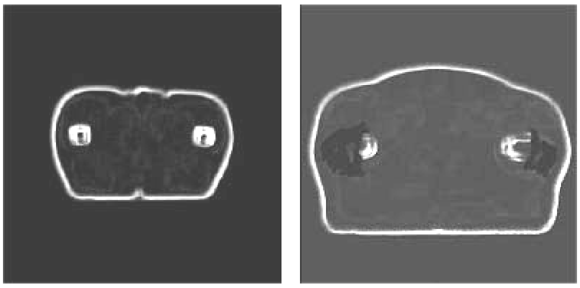

Figure 11. Responses of the MR images.

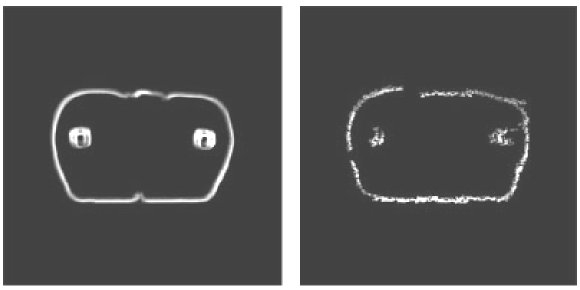

Figure 12. Result of the piece-wise mapping of the template (right) onto the reference image (left).

The results of the computational experiments presented in this section validate the proposed approach in the form of Image local response and its applicability to solving complex image mapping problems.

4. CONCLUSION

The paper proposes a novel image transformation in the form of Image Local Response (ILR) that can be used for solving complex image mapping problems. The proposed transformation brings together two approaches based on the pixel value distribution and image features.

Image local response is defined as the average value of the difference between the transformed and the original copies of the same image. Here, the transformation is small, i.e., the components of the corresponding parameter vector have sufficiently small unit values.

The response has a few interesting properties useful in image mapping:

- it significantly reduces the amount of information that has to be processed during the search for the correct mapping parameters,
- it retains the main features of the object shape, its contour,
- the algorithm for computing response values is inherently parallel,
- response provides a bell-shaped fitness landscape very well suited for solving image mapping problem with the evolutionary search,
- the ILR model can be used to effectively control and accelerate the search for the proper mapping.

The validity of the proposed image transformation is shown on complex image mapping problems formulated as multi-objective piece-wise imaging optimization problem.

5. REFERENCES

[Ash06a] Ashlock, D. Evolutionary computation for modeling and optimization. Springer, 2006.

[Bar89a] Barton, G. Elements of Green’s functions and propagation: Potentials, diffusion, and waves. Clarendon Press, Oxford, 1989.

[Bro98a] Brooks, R.R., and Iyengar, S.S. Multi-sensor fusion: Fundamentals and applications with software. Prentice Hall, New York, 1998.

[Ger02] Gere, J.M., and Timoshenko, S.P. Mechanics of materials. Nelson Thornes, 2002.

[Ima06a] Image processing toolbox user’s guide. The MathWorks, 2006.

[Mas05a] Maslov, I.V., and Gertner, I.. Reducing the computational cost of local search in the hybrid evolutionary algorithm with application to electronic imaging. Engineering Optimization 37, No 1, pp. 103-119, 2005.

[Mas08b] Maslov, I.V. Improving the performance of Evolutionary algorithms in imaging optimization. Ph.D. Thesis, City University of New York, New York, 2008.

[Muc98a] Muchnik, I., and Mottl, V. Bellman functions on trees for segmentation, generalized smoothing, matching and multi-alignment in massive data sets. DIMACS Technical Report 98-15, February 1998.

[Pit00a] Pitas, I. Digital image processing algorithms and applications. John Wiley & Sons, 2000.

[Ros76a] Rosenfeld, A., and Kak, A.C. Digital picture processing. Academic Press, New York, 1976.

[Seu00a] Seul, M., O’Gorman, L., and Sammon, M.J. Practical algorithms for image analysis: Description, examples, and code. Cambridge University Press, Cambridge, UK, 2000.