Landmark-based Non-rigid Registration via Graph Cuts

Herve Lombaert[1], Yiyong Sun[2], and Farida Cheriet[1]

[1] École Polytechnique de Montréal, Montréal Québec, Canada
[2] Siemens Corporate Research, Princeton NJ 08540, USA

Abstract. This paper presents an approach based on graph cuts initially used for motion segmentation that is being applied to the non-rigid registration problem. The main contribution of our method is the formulation of landmarks in the graph cut minimization framework. In the graph cut method, we add a penalty cost based on landmarks to the data energy. In the presence of a landmark, we adjust the *T-link* weights to cut strategic links. Our formulation also allows the spread of a landmark influence to its neighborhood. We first show with synthetic images that minimization with graph cuts can indeed be used for non-rigid registration and show how landmarks can guide the minimization process towards a customized solution. We later use this method with real images and show how landmarks can successfully guide the registration of a coronary angiogram.

1 Introduction

In image processing, many applications rely on a point by point correspondence for image analysis. In the medical field, comparing multiple images is an important task for diagnostics. For instance, when dealing with repeated imaging, pathologies are easier to monitor when they are correctly aligned on the images. The need for a common reference is essential to integrate complementary information acquired from different modalities (such as magnetic resonance or computed tomography) in the same visualization framework. The correct alignment of these images requires the transformation mapping all points from one image to the other image. Finding this transformation map is known as the registration process.

Non-rigid registration is a challenging problem as each point can have its own motion. Many approaches exist to solve the non-rigid problem ([1–3] for surveys). Popular methods involve parametric models such as free form deformations ([4]) or thin plate splines ([5]). These methods manipulate control points which are handled automatically or are inputs from the user. Applying forces to the deformation field is another popular approach. In [6] a physical model is used by adding an elasticity constraint to the deformation field. Horn and Schunk [7], and Lucas and Kanade [8], find the vector field by solving the optical flow equation which assumes a point keeps a similar intensity over time and the vector field remains smooth.

Optical flow is a seminal method for a very related problem – motion segmentation – where objects in images are found by segmenting the motion field. Among the approaches proposed to solve motion segmentation, energy minimization via graph cuts ([9–12]) finds a solution which has the nice property of being within a known factor of the global minimum ([12]). Further work with graph cuts ([13, 14]) handle occlusions in visual correspondence and motion segmentation. However, when the deformation field is assumed to be smooth and to have no break, occlusions can be negligible. In non-rigid registration, this happens when all points of an object have a correspondence on the other image.

The graph cut minimization framework as described in [12] can be used to solve the non-rigid, unsupervised, registration problem. It has been used in [15] to find visual correspondances between two images. We introduce in this paper a formulation for landmarks that guides the minimization towards a customized solution. This helps solve ambiguities and allows a supervision of the registration process. The method will first be explained. Later, landmarks are being studied with experiments using synthetic and real images.

2 Method

The main objective of this paper is to introduce landmarks in the graph cut minimization framework. We propose a new energy formulation to solve a landmark based non-rigid registration. First, we formulate non-rigid registration as an energy minimization problem that can be solved via graph cuts. Second, we introduce landmark in the minimization framework.

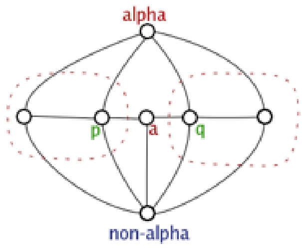
2.1 Non-rigid Registration as an Energy Minimization

In a non-rigid deformation, although each pixel of an image can move freely, the motion is assumed to be locally coherent. The deformation field undergoes two forces, one that matches the warped image with the original image, the second that keeps the deformation field smooth.

Let the deformation field be represented by a vector field f where each vector has a deformation $f_p \in \mathcal{L}$ defined in $\mathbb{R}^D$. The deformation field between two images I_1 and I_2 can be recovered by minimizing the following energy:

$$E(f) = \sum_{p \in I} D\left(I_1(p), I_2(p + f_p)\right) + \lambda \sum_{p,q \in \mathcal{N}} V(f_p, f_q) \tag{1}$$

The first term, $D_p(f_p)$, measures how the data differs between the original image and the warped image. For instance, it can be the squared sum of differences. When using multimodal images, mutual information can be used ([16]). The second term, $V_{p,q}$, measures the smoothness of the deformation field. It can be the norm of the difference of two neighboring vectors. The parameter λ controls the smoothness.



$w(\alpha, p) = D(\alpha)$
$w(\bar{\alpha}, p) = D(f_p)$, if $p \notin \mathcal{P}_\alpha$
$w(\bar{\alpha}, p) = \infty$, if $p \in \mathcal{P}_\alpha$
$w(p, a) = V(f_p, \alpha)$, and similar for $w(a, q)$
$w(p, \bar{\alpha}) = V(f_p, f_q)$
$w(p, q) = V(f_p, \alpha)$, when $f_p = f_q$
$p \in \mathcal{P}_\alpha$ if the link $(p - \alpha)$ is cut
$p \in \mathcal{P}_{f_p}$ if the link $(p - \bar{\alpha})$ is cut

Fig. 1. An example of a graph used in an iteration of the α expansion algorithm. Here, four variables are implemented with four graph nodes. Two special nodes α and $\bar{\alpha}$, connected to the variable nodes with data links, optimally assign the value α or f_p to each variable. All link weights w are provided on the right.

Boykov *et al* ([12]) show two algorithms based on graph cuts: the α expansion and the $\alpha - \beta$ swap algorithms. They minimize any functional (1) whose smoothness term is either metric or semi-metric, that is:

$$V(\alpha, \beta) = 0 \Leftrightarrow \alpha = \beta$$
$$V(\alpha, \beta) = V(\beta, \alpha) \geq 0$$
$$V(\alpha, \beta) \geq V(\alpha, \gamma) + V(\gamma, \beta)$$

If the first two conditions are satisfied, the smoothness term is said to be semi-metric, and the $\alpha - \beta$ swap algorithm can be used. However, if all three conditions are satisfied, the term is said to be metric, and the α expansion can also be used. A metric term is chosen for (1), for instance:

$$E(f) = \sum_{p \in I} (I_1(p) - I_2(p + f_p))^2 + \lambda \sum_{p,q \in \mathcal{N}} ||f_p - f_q|| \tag{2}$$

The main idea of the α expansion algorithm is to iteratively minimize the energy, testing one value α at a time. Here, α can be any allowed deformation, e.g., $\alpha = (+1, +1)$ pixels. At each step, graph cuts (Fig. 1) optimally assign $f_p = \alpha$ to the deformation field f. The α region is said to expand. In [12], a more in depth explanation of the α expansion algorithm is provided.

One attractive property of the α expansion algorithm is that it guarantees a convergence towards a local minimum $E(\hat{f})$ that is within a known factor of the global minimum $E(f^*)$:

$$E(\hat{f}) \leq 2 \left(\frac{\max_{f_p \neq f_q} ||f_p - f_q||}{\min_{f_p \neq f_q} ||f_p - f_q||} \right) E(f^*)$$

2.2 Landmarks in the Graph Cut Framework

The recovery of the deformation field as just proposed earlier is a non supervised process. The inputs are the two images, and the minimization finds the

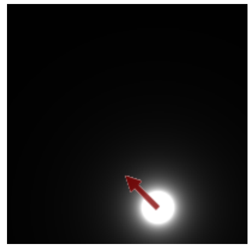


Fig. 2. Penalty associated with a landmark. The landmark transformation ($f(k) = \{-25, +25\}$) is displayed with the arrow – The image shows for each pixel the cost to assign a transformation f_p such that $f_p \neq f(k)$, i.e., different than the landmark transformation. The penalty cost gets high in the landmark neighborhood and tends to zero when it is far from the landmark. This cost will eventually change the data cost D_p for each pixel.

best deformation field. Introducing landmarks into this framework would allow interaction and control from the user.

A landmark $k \in \mathcal{K}$ defines a motion $f(k) \in \mathcal{L}$ at its position $p(k) \in I$. In other words, the pixel siting at a landmark position $p(k)$ on the original image is known to have moved by $f(k)$ on the second image. This can be seen as a hard constraint on the motion field.

In the minimization framework, that means, where there is a landmark k, the transformation of pixel p should always be the label $f(k)$. In the graph cut methods, wherever there is a landmark k at a pixel p, it should be impossible to cut the data link, or *T-link*, $(\alpha - p)$ with $\alpha \neq f(k)$ (i.e. the pixel p cannot be assigned something else than the label $f(k)$). This is enforced by setting these link weights to infinity, or more generally:

$$D_p(\alpha) = \infty \text{ if } \alpha \neq f(k) \text{ at a landmark site } k$$

A landmark can influence its neighborhood by using a penalty cost over the whole image. This penalty is infinity at the landmark site and quickly decreases around the landmark. For instance, the data term can be:

$$D_p(\alpha) = d_p(\alpha) + \mu \sum_{\substack{k \in \mathcal{K} \\ f(k) \neq \alpha}} \frac{r^2}{|p(k) - p|^2} \tag{3}$$

Besides the standard data cost $d_p(\alpha)$ (e.g. sum of squared differences), all landmarks, whose deformation $f(k) \neq \alpha$, provide a penalty based on the distance of the point p to each landmark (Fig 2). Here r is the distance within which the landmark k has a strong effect on point p. Beyond this distance, the penalty drops below 1 and tends to 0 at infinity. If the point p is at a landmark site, the penalty is infinity and it is not possible to assign the value α at this point. μ is a factor controlling the penalty of landmarks.

Note here that the cost to assign α does not depend on landmarks whose deformation is $f(k) = \alpha$. Regions with the label $f(k)$ could not only happen around the landmark site, but anywhere. By adding no penalty for $f(k) = \alpha$, all regions are given an equal opportunity to be assigned $f(k)$.

3 Results

To validate our method we use three experiments. The first shows how to recover a non-rigid deformation field. The second shows how landmarks can influence the deformation map. The third is a real case example with a coronary angiogram. There, we show that the use of landmarks can improve the visual quality of the deformation map in a graph cut minimization framework.

During the experiments, the data term of (2) has been replaced with the landmark formulation (3) (second term, below). This yields to:

$$E(f) = \sum_{p \in I} (I_1(p) - I_2(p + f_p))^2 + \mu \sum_{p \in I} \sum_{\substack{k \in \mathcal{K} \\ f(k) \neq f_p}} \frac{r^2}{|p(k) - p|^2} + \lambda \sum_{p,q \in \mathcal{N}} ||f_p - f_q||$$

3.1 Synthetic Transformation

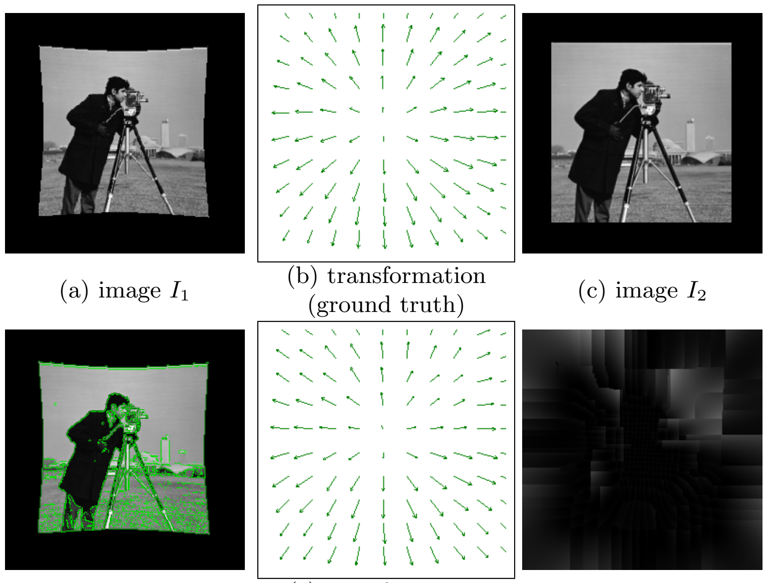
In this experiment a recovered deformation map is compared with its ground truth. The figure 3b shows the known transformation between image I_1 (Fig 3a) and image I_2 (Fig 3c) such that for a pixel p, $I_1(p) = I_2(p + f_p)$.

It is known from the ground truth that a pixel moves at most by 15 pixels. The used label set is then $f_p \in \{(dx, dy)\}$, where $dx \in [-15; +15]$ and $dy = [-15; +15]$. A smoothness coefficient of $\lambda = 0.01$ has been used. Minimization with graph cuts could successfully recover the deformation map (Fig 3e). This transformation has been applied on image I_2 which resulted in an image I_2' (Fig 3d) very close to I_1. The registered image I_2' is constructed from pixels of I_2 such that $I_2'(p) = I_2(p + f_p)$.

The recovered map is almost identical to the ground truth. The exaggerated difference (Fig 3f) shows it is almost zero. However disconcordancies appear at the top right corner of the image (gray sky). This region is mainly uniform and many transformations can lead to a correct solution. In such regions, the optimization tends to produce piecewise uniform rigid transformations.

3.2 Influence of Landmarks

The registration process is an ill-posed problem. Many transformations can lead to the given images. Here, we show that landmarks can favor one solution to another. Two vertical bars are moved apart by 20 pixels (original on Fig 4a, final on Fig 4b). Many transformations can move the two bars apart. Three labels are used $f_p \in \{(-20; 0), (+20; 0), (0, 0)\}$. A pixel can move left, right,



(a) image I_1 (b) transformation (ground truth) (c) image I_2

(d) registerd I_2 with I_1 (e) transformation (recovered) (f) difference

Fig. 3. Recovering a non-rigid transformation map: (a,c) Image I_1 and I_2, and (b) the ground truth f^* such that $I_1(p) = I_2(p + f^*_p)$. (d) shows the registered image $I_2(p + \hat{f})$ overlayed with I_1 in light contours. (e) shows the recovered transformation map $\hat{f}$. (f) shows the magnitude of the squared difference of the deformation map $\hat{f}$ with its ground truth f^* (intensities are multiplied by 5 for a better comparison).

or remain static. When running the graph cut optimizer, the deformation field shown on the figure 4c shows a line delimiting the motions moving the vertical bars apart.

As the image intensities are either black or white, this line could be anywhere between the two vertical bars. By placing two landmarks (Fig 4a), the motion of white regions is fixed to the left on top of the image, and to the right on its bottom. The smoothness coefficient has been set to $\lambda = 0.01$ and the landmark influence to $\mu = 1$ and $r = 3$. The deformation field recovered by the graph cut optimizer shows that the landmarks dragged correctly the white regions (Fig 4d,e).

With graph cuts, the graph topology is known to change the cut metric. With a neighborhood of 4 pixels, A cut between two neighbors can either be vertical or horizontal. That means a graph cut such as on the figure 4d is cheaper than the cut on the figure 4e. With a neighborhood of 8 pixels, diagonal cuts are now permitted and the length of the graph cut on figure 4e becomes cheaper.

[1] each landmark is displayed with a small arrow. Its base is where a pixel was on the first image, and its tip points to where the pixel moved on the second image.

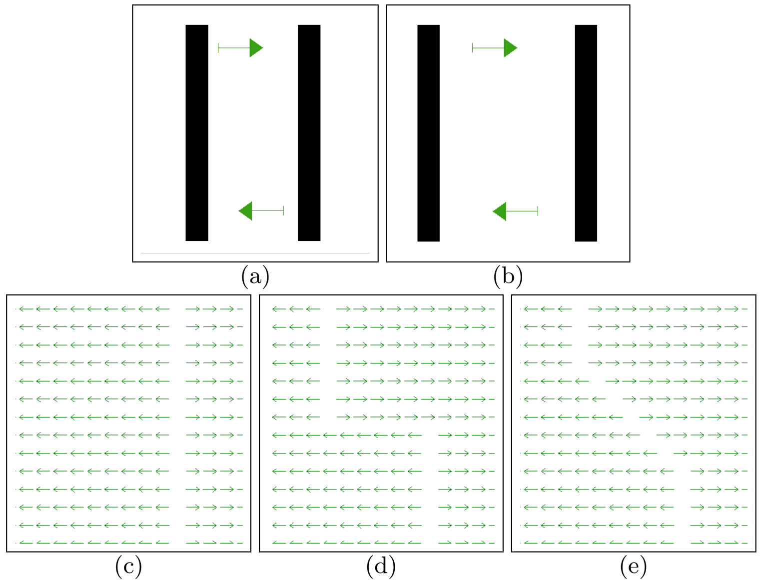
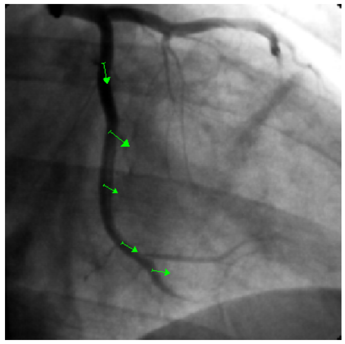


Fig. 4. Landmark influence: (a) original image with control points[1], (b) final image. (c) shows the recovered transformation map without landmarks. (d) shows the deformation map with the use of landmarks and a 4 neighborhood system, and (e) with a 8 neighborhood system.

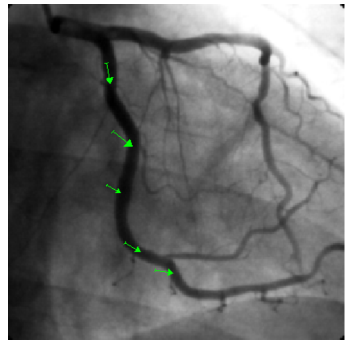
3.3 Real Case

Knowing that graph cuts can recover a non-rigid deformation field and that landmarks can influence the solution, we applied this method on a real case image. Two arbitrary frames from a coronary cineangiogram have been used. From the images, the general motion of the vessels is downward. Heuristic inspections yield that the vessel can move horizontally with $dx \in [-18; +18]$ pixels, and vertically with $dy \in [-18; +5]$ pixels. This is used to reduce the search domain and speed up the optimization. The smoothness coefficient is set to $\lambda = 0.01$ and the landmark influence to $\mu = 1$ and $r = 3$.

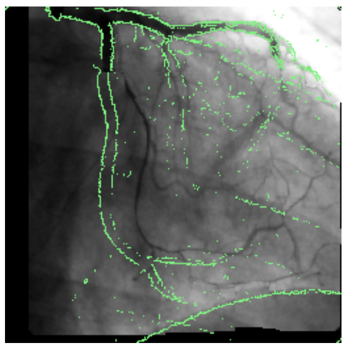
On the figure 5c,d, no landmarks are used. The minimization prefered to register the dark spine on the left side of the image rather than register the vessels. That is because the vessels are too thin compared to the background and as the graph cut minimization finds a solution close to the global optimum, the recovered deformation field favored the registration of the background. This can be solved by using landmarks located on the vessels (Fig 5e,f). The deformation field becomes constrained in the vessels area and the solution on the figure 5e shows an improvement on their alignment.



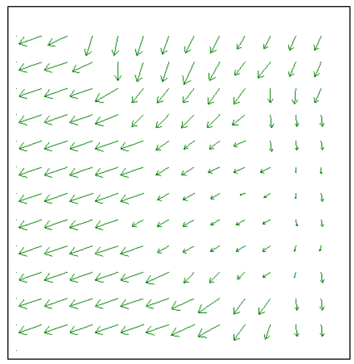
(a) first frame I_1 with landmarks



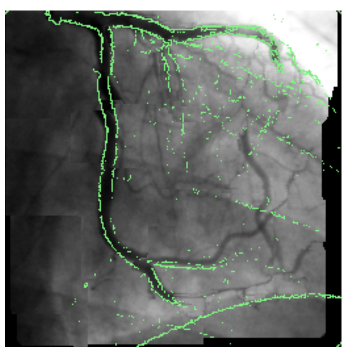
(b) second frame I_2



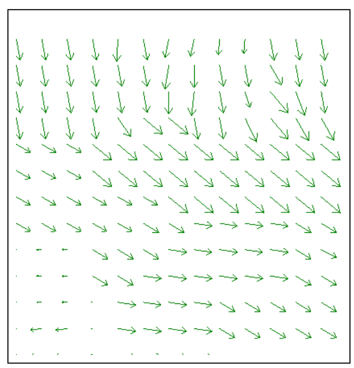
(c) registered I_2 without landmarks



(d) deformation field



(e) registered I_2 using landmarks



(f) deformation field

Fig. 5. Non-rigid registration with landmarks: The first row shows two frames from a coronary angiogram with the landmarks indicated with small arrows; the second row shows the registration without landmarks, as seen on (c) the minimization prefers to register the background (dark spine on the left and diaphragm on the bottom) rather than the vessels (which are much smaller compared to the background); the third row shows the registration with landmarks, as seen on (e) landmarks help register the vessels. Images on (c) and (e) show the registered image I_2 with the original frame I_1 (in light contours). Images on (d) and (f) show the deformation map.

4 Discussion

In this paper we proposed to use landmarks with the graph cut minimization framework for non-rigid registration. The original framework has been used for many applications including segmentation, restoration of images, visual correspondence and in particular motion segmentation. As this last problem is closely related to image registration, it is possible to use graph cuts to recover a non-rigid transformation. Furthermore, a formulation for landmarks in this framework can guide the minimization process towards a solution influenced externally. This way, the registration process becomes supervised and interactive. The results showed an example where a physician can use landmarks to correct an unsupervised registration. Landmarks could also be set by an automatic process, which could rely on a segmentation result. Furthermore the concept of landmarks could be applied to other graph cut applications. They can set known visual correspondences, or known intensities in image restoration.

The combinatorial nature of the graph cut method only allows a finite set of possible transformations. The label set must include all possible transformations. This is a limitation for the use of graph cuts in non-rigid registration. A different approach using continuous values for transformations as well as a better way to control the smoothness of the deformation field will allow a more complete framework for non-rigid registration with graph cuts. Future work will aim into these issues.

Acknowledgement

The authors wish to acknowledge the financial support from Siemens Corporate Research (SCR) and the Natural Sciences and Engineering Research Council of Canada (NSERC), as well as Dr. Chenyang Xu, for helpful discussions, and Dr. Frank Sauer at SCR.

References

1. W. R. Crum, T. Hartkens, and D. L. Hill. Non-rigid image registration: theory and practice. *The British Journal of Radiology*, 77(2):140–153, 2004.
2. J. B. Maintz and M. A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, March 1998.
3. T. McInerney and D. Terzopoulos. Deformable models in medical image analysis: a survey. *Medical Image Analysis*, 1(2):91–108, June 1996.
4. D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, August 1999.
5. F. L. Bookstein. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.
6. G. E. Christensen, M. I. Miller, M. W. Vannier, and U. Grenander. Individualizing neuroanatomical atlases using a massively parallel computer. *Computer*, 29(1):32–38, January 1996.

7. B. Horn and B. Schunck. Determining optical flow. *Artificial Intelligence*, 17(1-3):185–203, August 1981.
8. B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proceedings of the Seventh International Joint Conference on Artifical Intelligence, IJCAI 1981*, pages 674–679, Vancouver, Canada, 1981.
9. Y. Boykov, O. Veksler, and R. Zabih. Markov random fields with efficient approximations. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR 1998*, pages 648–655, Santa Barbara, CA, June 1998. IEEE Computer Society.
10. Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. In *Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 359–374, 2001.
11. Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9):1124–1137, September 2004.
12. Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1222–1239, 2001.
13. V. Kolmogorov and R. Zabih. Computing visual correspondence with occlusions via graph cuts. In *Proceedings of the Eighth IEEE International Conference on Computer Vision, ICCV 2001*, volume 2, pages 508–515, Vancouver, Canada, July 2001.
14. M. Shah and J. Xiao. Motion layer extraction in the presence of occlusion using graph cuts. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 27(10):1644–1659, October 2005.
15. P. Bhat, K. C. Zheng, N. Snavely, A. Agarwala, M. Agrawala, M. F. Cohen, and B. Curless. Piecewise image registration in the presence of multiple large motions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern, CVPR 2006*, volume 2, pages 2491–2497, New York, NY, June 2006. IEEE Computer Society.
16. J. Kim, V. Kolmogorov, and R. Zabih. Visual correspondence using energy minimization and mutual information. In *Proceedings of the Ninth IEEE International Conference on Computer Vision, ICCV 2003*, volume 2, pages 1033–1040, Nice, France, October 2003.