A Study on Image Change Detection Methods for Multiple Images of the Same Scene Acquired by a Mobile Camera

Guntur Tanjung

School of Mechanical Engineering

The University of Adelaide

South Australia 5005

Australia

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in Mechanical Engineering on the 12th October 2009

APPENDIX 1

In this study, the Scale Invariant Feature Transform (SIFT) operator [Lowe, 2004] was used to automatically extract and match local features from both reference and input images. The local features were extracted by the SIFT operator from three templates which were chosen and cropped manually off-line in advance from the reference image. The local features provided by the SIFT operator were invariant to rotation, scaling, and slightly tolerant of small changes in illumination.

The SIFT operator. The Scale Invariant Feature Transform was developed by David Lowe (Lowe, 2004) to find stable image features that can be used for object recognition. These features are invariant to rotation and scaling, as well as being relatively tolerant of small changes in illumination. This kind of invariance means that these same features are also likely to be detected in images where the object has been rotated or is viewed from a slightly different perspective or distance. There are four stages to the SIFT algorithm.

- Scale-space extrema detection
- Keypoint localization
- Orientation assignment
- Keypoint descriptor

Scale-space extrema detection. In this initial stage of the algorithm, a source image I(x, y) is convoluted using a 2-dimensional Gaussian function G(x, y, σ) to build an image pyramid. This creates a 3-dimensional representation with x and y representing the 2 image axis, and σ representing the scale axis of the image. Increasing σ moves the image plane upward in scale-space, and the image becomes increasingly more blurred as the higher spatial frequencies are filtered out.

The function G(x, y, σ) is defined as follows:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \cdot e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad \text{.......... ....}(1)$$

Convoluting the image I(x, y) with G(x, y, σ) creates the scale space representation L(x, y, σ) of the image.

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(2)}$$

The SIFT algorithm uses extrema found in the difference-of-Gaussian function from the convolved image pyramid that is computed from two adjacent scales in scale-space separated by a constant factor k.

The difference-of-Gaussian function D(x, y, σ) of the convoluted image L(x, y, σ) is the subtraction of two adjacent scales in the Gaussian scale-space pyramid separated by constant factor k.

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(3)}$$

To create the image pyramid, discrete intervals in scale-space are sampled by increasing the scale parameter σ by a constant amount. In Lowe's (2004) paper, the image is down sampled by a factor of 2 when the scale parameter σ doubles so that an octave has half the previous octave's dimensions. This provides a great increase in processing speed with a negligible loss in accuracy.

Each of the octaves are split in to s intervals, and the factor k required to provide the correct number of intervals over an octave can be calculated using 21/s. To search a complete octave for extrema detection s+3 intervals are required per octave for s+2 difference-of-Gaussian intervals. For instance, if s = 2 then the number of Gaussian pyramid intervals per octave that would need to be created would be 2 + 3 = 5 to create 2 + 2 = 4 DoG intervals to search for extrema cover a complete octave. Of these DoG intervals, only 2 levels will be checked against the level above and below (for DoG intervals 1 and 4 there are no intervals to check below and above respectively, so they are only used during the search over intervals 2 and 3).

Extrema are found in the DoG intervals by comparing a sample to it's 8 neighbours on the same level, 9 in the interval above and 9 in the interval below. If and only if the sample is either greater than or less than all of its 28 neighbours then it is considered as a potential keypoint.

Keypoint localization. Once all potential keypoint candidates have been identified, the keypoints need to be checked for stability. The first test that needs to be made is that of contrast. If the value of D(x, y, σ) at the keypoint location is less than a contrast threshold constant, then it is discarded as unstable and susceptible to low levels of noise. A value of |D(x, y, σ)| less than 0.03 is used to filter out low contrast keypoints in Lowe's (2004) paper.

The most recent SIFT paper proposes a method of interpolating a keypoint by fitting a 3D parabola to the nearby sample data. If the calculated offset from the sample point is greater than 0.5 in any dimension, then the keypoint is moved to that position instead. Lowe (2004) proposes that this improves the stability of the keypoints found in the image.

The interpolated position is found by solving the following equation to find the offset:

$$\hat{x} = -\frac{\delta^2 D}{\delta x^2} \cdot \frac{\delta D}{\delta x} \quad \text{....(4)}$$

Where $\hat{x}$ is the offset vector from the sample point, and D and its derivatives are evaluated using local pixel differences around the sample point.

Lowe (2004) proposes that the function value at this final offset location D(ˆx) is useful for rejecting unstable points with low contrast. This value can be found by evaluating:

$$D(\hat{x}) = D + \frac{1}{2} \cdot \frac{\delta D^T}{\delta x} \hat{x} \quad \text{....(5)}$$

Next keypoints that are situated along edges need to be rejected as they are poorly defined and are likely to be susceptible to small amounts of noise (Lowe 2004). Keypoints will have a small curvature along and edge, and a large curvature across the edge. These points can be rejected by finding the ratio of the principle curvatures at the sample point and rejecting ratios that are too large.

These principle curvatures can be found by calculating the Hessian matrix H at the keypoint location, using differences of neighbouring pixel samples to find the derivatives. Note that the matrix is symmetric so Dxy = Dyx.

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad \text{.........(6)}$$

The ratio of the principle curvatures can be found by calculating the trace T r(H) and determinant Det(H) of the Hessian matrix:

$$Tr(H) = D_{xx} + D_{yy} \quad \text{.........(7)}$$

$$Det(H) = D_{xx} \cdot D_{yy} - D_{xy}^{2} \quad \text{.........(8)}$$

Then the following equation is evaluated to find the ratio of principle curvatures and check that it is below a desired threshold value r (Lowe 2004).

$$\frac{Tr(H)^2}{Det(H)} < \frac{(r+1)^2}{r} \quad \text{.........(9)}$$

Keypoints that have a curvature threshold less than r are rejected, as well as points that have a negative value for Det(H). Keypoints with ratios greater than r = 10 were rejected in Lowe (2004).

Orientation assignment. Once unstable keypoints have been discarded, orientations need to be assigned to each of the remaining keypoints so that local pixel data can be described relative to the orientation of a keypoint for rotational invariance. The image gradient orientations and magnitudes of all the sample pixels on the same level as the keypoint that are within a Gaussian window of 1.5 times the scale of the keypoint are used to calculate a histogram. The histogram consists of 36 bins (one for each 10◦ step). Each sample contributes to the appropriate bin by its magnitude weighted by the Gaussian window. For efficiency the orientations and magnitudes of the pixel data are pre-calculated during the creation of the image pyramid.

Let the finite central differences across x and y at pixel location $(x, y)^T$ be $\delta_x(x, y)$ and $\delta_y(x, y)$ respectively, then:

$$\delta_x(x, y) = L(x + 1, y) - L(x - 1, y) \quad \text{……………………………….(10)}$$

$$\delta_y(x, y) = L(x, y + 1) - L(x, y - 1) \quad \text{……………………………….(11)}$$

The magnitude m(x, y) at a sample pixel location can be found by evaluating the following equation:

$$m(x, y) = \sqrt{\delta_x(x, y)^2 + \delta_y(x, y)^2} \quad \text{..(12)}$$

The orientation θ(x, y) of the gradient is relative to the image space x axis, and can be determined as follows:

$$\theta(x, y) = \tan^{-1}\left(\frac{\delta_y(x, y)}{\delta_x(x, y)}\right) \text{...(13)}$$

Once the histogram has been calculated, keypoints are created for each orientation that has a value of 80% of the maximum histogram value or more. Lowe (2004) states that this contributes significantly for stability of matching. For each of these peak orientations, a parabola is fit to the values of the 3 nearest bins to interpolate the orientation for better accuracy.

Keypoint descriptor. Once all keypoint locations have been determined and have orientations assigned to them, the next stage is to create a descriptor to represent the image data around the keypoint in an invariant form. Keypoint descriptors as used in Lowe (2004) are composed of a 4x4 grid of histograms formed from 4x4 pixel subregions from a larger 16x16 sample array. Histograms consist of 8 bins, one for each 45◦ step. The magnitudes of each of the sample points are weighted by a Gaussian window of width of half the keypoint's scale. To avoid boundary effects, a sample contributes to more than one bin in the histogram weighted by a factor of $1 - d$ where d is distance in histogram step units from a bin orientation. A sample also contributes to bins in adjacent histograms in the descriptor, again weighted by $1-d$ where d is distance in descriptor histogram spacing units (4 for a 16x16 sample block).

Reference

Lowe, D. G., 2004, "Distinctive image features from scale-invariant keypoints." *International Journal of Computer Vision* 60(2): 99-110.

References

Aach, T., and Kaup, A., 1995, "Bayesian algorithms for adaptive change detection in image sequences using Markov random fields", *Signal Processing: Image Communication*, vol. 7, pp. 147-160 *(Cited on page 16)*

Amit, Y., and Kong, A., 1996, "Graphical templates for model registration", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 3, pp. 225-236 *(Cited on page 9)*

Arques, P., Aznar, F., Pujol, M., and Rizo, R., 2006, "*Real Time Image Segmentation Using an Adaptive Thresholding Approach",* Book Chapter, Lecture Notes in Computer Science, pp. 389-398, Springer Berlin / Heidelberg *(Cited on page 117)*

Asai, K., and Aschmann, C., 1995, "*Fuzzy Systems for Information Processing*", IOS Press, Amsterdam *(Cited on page 97)*

Baldock, R. and Graham, J., 2000, "*Image Processing and Analysis: A Practical Approach"*, Oxford University Press, New York *(Cited on page 116)*

Basu, M., 2002, "Gaussian-based edge-detection methods-a survey", *IEEE Trans. On Systems, Man and Cybernetics*, vol. 32, Issue: 3, pp. 252-260 *(Cited on pages 114 and 115)*

Benabdelkader, C., Burlina, P., and Davis, L., 2000, "Single Camera Multiplexing for Multi-Target Tracking", *In Multimedia Video-based Surveillance System*, pp. 130-142, G.L. Foresti, P. Mahonen, and C. Regazzoni editor, Kluwer Academic Publisher, Norwell *(Cited on page 21)*

Birchfield, S., and Tomasi, C., 1999, "Multiway Cut for Stereo and Motion with Slanted Surfaces", *Proc. Int'l Conf. Computer Vision*, vol. 1, pp. 489-495 *(Cited on pages 14 and 23)*

Bosc, M., Heitz, F., Armspach, J. P., Namer, I., Gounot, D., and Rumbach, L., 2003, "Automatic change detection in multimodal serial MRI: application to multiple sclerosis lesion evolution", *Neuroimage*, vol. 20, pp. 643–656 *(Cited on pages 8)*

Boykov, Y., Veksler, O., and Zabih, R., 1998, "Markov Random Fields with Efficient Approximations", *Proc. Computer Vision and Pattern Recognition*, pp. 648-655 *(Cited on page 14)*

Bruzzone, L., and Prieto, D. F., 2002, "An adaptive semiparametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images", *IEEE Trans. Image Processing*, vol. 11, no. 4, pp. 452–466 *(Cited on page 8)*

Cavallaro, A., and Ebrahimi, T., 2002, "Accurate video object segmentation through change detection", in *Proc. of IEEE International Conference on Multimedia and Expo*, pp. 445–448 *(Cited on page 19)*

Chalermwat, P., and El-Chazawi, T. A., 1999, "Multiresolution image registration using genetics", *ICIP*(2), pp. 452-456 *(Cited on page 12)*

Collins, J. B., and Woodcock, C. E., 1996, "An assessment of several linear change detection techniques for mapping forestmortality using multitemporal Landsat TM data", *Remote Sensing Environment*, vol. 56, pp. 66–77 *(Cited on page 8)*

Collins, R., Lipton, A., and Kanade, T., 2000, "Introduction to the special section on video surveillance", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, no. 8, pp. 745–746 *(Cited on page 8)*

Dai, X., Khorram, S., and Cheshire, H., 1996, "Automated image registration for change detection from thematic mapped imagery", *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*, Lincoln, Nebraska, vol. 111, pp. 1609-1611 *(Cited on page 9)*

Dhond, U. R. and Aggarwal, J. K., 1989, "Structure from Stereo – A Review", *IEEE Trans. Syst., Man and Cybern.* vol. 19, pp. 1489-1510 *(Cited on page 13)*

Durucan, E., and Ebrahimi, T., 2001, "Change detection and background extraction by linear algebra", *Proceeding of the IEEE*, vol. 89, no. 10, pp. 1368-1381 *(Cited on page 17 and 18)*

Edgington, D., Dirk, W., Salamy, K., Koch, C., Risi, M., and Sherlock, R., 2003, "Automated event detection in underwater video", in *Proc. MTS/IEEE Oceans 2003 Conference* *(Cited on page 8)*

Fang, C.-Y., Chen, S.-W., and Fuh, C.-S., 2003, "Automatic change detection of driving environments in a vision-based driver assistance system", *IEEE Trans. Neural Networks*, vol. 14, no. 3, pp. 646–657 *(Cited on page 8)*

Fonesca, L.M, and Costa, M.H.M, 1997, "Automatic registration of satellite images", *Proceedings on IEEE Transactions of Computer Society*, pp. 219-226 *(Cited on page 11)*

Fonseca, L. M. G., and Manjunath, B. S., 1996, "Registration techniques for multisensor remotely sensed imagery", *Photogrammetric Engineering & Remote Sensing*, vol. 62, no. 9, pp. 1049-1056 *(Cited on page 45)*

Francois, A. R., and Medioni, G. G., 1999, "Adaptive color background modeling for real-time segmentation of video streams", in *Proceedings of the*

International Conference on Imaging Science, Systems, and Technology, pp. 227–232, (Las Vegas, NA) *(Cited on page 18)*

Fua, P., 1993, "A parallel stereo algorithm that produces dense depth maps and preserves image features", *Machine Vision and Applications*, vol. 6, pp. 35-49 *(Cited on page 74)*

Goldberg, D. E., 1989, "Genetic algorithm in search: optimization and machine learning", Reading, Mass. Addison-Wesley *(Cited on page 12)*

Gonzalez, R. C., and Woods, R. E., 2008, "Digital Image Processing," 3rd edition. Prentice Hall, Upper Saddle River, New Jersey *(Cited on pages 78, 129, 132 and 134)*

Goshtasby, A. A., 2005, "2-D and 3-D image registration for medical, remote sensing, and industrial applications", John Wiley & Sons, New Jersey *(Cited on pages 9, 22 and 63)*

Haering, N., Venetianer, P. L., and Lapton, A., 2008, "The evolution of video surveillance: an overview", *Machine Vision and Applications*, vol. 19, pp. 279-290 *(Cited on page 3)*

Haritaoglu, D., Harwood, I., and Davis, L., 1998, "W4: Who? when? where? what? a real time system for detecting and tracking people", in *International Conference on Face and Gesture Recognition*, 3, ed. *(Cited on page 18)*

Hsu, Y. Z., 1984, "New likelihood test methods for change detection in image sequences", *Computer Vision Graph. Image Process.*, vol. 26, pp. 73-106 *(Cited on page 16)*

Hu, Z., and Acton, S.T., 2000, "Morphological pyramid image registration", *IEEE south west sysposium,* p. 227 *(Cited on page 12)*

Huertas, A., and Nevatia, R., 2000, "Detecting changes in aerial views of man-made structures", *Image and Vision Computing*, vol. 18, no. 8, pp. 583–596 *(Cited on page 8)*

Hong, H., and Woo, W., 2003, "A background subtraction for a vision-based user interface", in *ICICS-PCM* *(Cited on page 19)*

Hong, T., and Schowengerdt, R. A., 2005, "A robust technique for precise registration of radar and optical satellite images", *Photogrammetric Engineering & Remote Sensing*, vol. 71, no. 5, pp. 585-593 *(Cited on page 9)*

Intille, S., and Bobick, A., 1994, "Incorporating intensity edges in the recovery of occlusion regions", *International Conference on Pattern Recognition, Jerusalem*, vol. I, pp. 674-677 *(Cited on page 73)*

Kang, Chung-Chia and Wang, Wen-June, 2007, "A novel edge detection method based on the maximizing objective function", *Pattern Recognition,* vol. 40, Issue 2, pp. 609-618 *(Cited on page 114)*

Keller, Y., and Averbuch, A., 2002, "FFT based image registration", *IEEE International Conference-ICASSP 2002*, Orlando *(Cited on page 11)*

Kettnaker, V., and Zabih, R. 1999, "Counting People from Multiple Cameras", *In Proceedings of the IEEE International Conference on Multimedia Computing and Systems*, pp. 267-271, Florence *(Cited on page 20)*

Lampropoulos, G.A., Chan, J., Secker, J., Li, Y., and Jouan, A., 2003, "Automatic registration of Electro-optical and SAR images", *Proceedings of IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, pp. 219-226 *(Cited on page 9)*

Lebart, K., Trucco, E., and Lane, D. M., 2000, "Real-time automatic sea-floor change detection from video", in *MTS/IEEE OCEANS 2000*, pp. 337–343 *(Cited on page 8)*

Lowe, D. G., 2004, "Distinctive image features from scale invariant keypoints", *International Journal of Computer Vision*, vol. 60, no. 2, pp. 99-110 *(Cited on pages 6, 26 and 46)*

McKenna, S., Jabri, S., Duric, Z., Rosenfeld, A., and Wechsler, H., 2000, "Tracking groups of people", *CVIU* 80, pp. 42–56 *(Cited on page 19)*

Nakamura, Y., Matsuura, T., Satoh, K., and Ohta, Y., 1996, "Occlusion detectable stereo --- Occlusion patterns in camera matrix", *IEEE Conf. on Computer Vision and Pattern Recognition* *(Cited on page 73)*

Oberti, F., Marcenaro, L., and Regazzoni, C. S., 2002, "Real-time change detection methods for video-surveillance systems with mobile camera"" *In XI European Signal Processing Conference 2002 (EUSIPCO'2002)*, Toulouse, France, pp. 37-40 *(Cited on page 21)*

Okutomi, M., and Kanade, T., 1992, "A Locally Adaptive Window for Signal Matching",. *Intl. Journal of Computer Vision*, vol.7, no.2, pp. 143-162 *(Cited on pages 15 and 45)*

Orduyilmaz, A., 2006, "Automated image registration and mosaicking for multi-sensor images acquired by a miniature unmanned aerial vehicle platform", Master Thesis, Mississippi State University *(Cited on page 22)*

Otsu, N., 1979, "A Threshold Selection Method from Gray-Level Histograms", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62-66 *(Cited on pages 103 and 104)*

Peihua, Q., 2007, "Jump surface estimation, edge detection, and image restoration", *Journal of the American Statistical Association*, vol. 102, num. 478, pp. 745-756 *(Cited on pages 23, 114 and 115)*

Phong, B. T., 1975, "Illumination for computer generated pictures", *Communication ACM*, vol. 18, pp. 311-317 *(Cited on pages 17)*

Primdahl, K., Katz, I, Feinstein, O., Mok, Y., Dahlkamp, H., Stavens, D., Montemerlo, M., and Thrun, S., 2005, "Change Detection from Multiple Camera Images Extended to Non-Stationary Cameras", *In Proceedings of Field and Service Robotics 2005 (FSR05)***,** Port Douglas, Australia *(Cited on pages 21)*

Radke, R. J., Andra, S., Al-Kofahi, O., and Roysam, B., 2005, "Image change detection algorithms: a systematic survey", *IEEE Transactions on Image Processing*, vol. 14, no. 3, pp. 1-14 *(Cited on pages 3 and 63)*

Reddy, B.S., and Chatterji, B.N, 1996, "An FFT-based technique for translation, rotation and scale invariant image registration", IEEE Transactions on Image Processing, vol. 5, no. 8, pp. 1266-1271 *(Cited on pages 10 and 11)*

Rey, D., Subsol, G., Delingette, H., and Ayache, N., 2002, "Automatic detection and segmentation of evolving processes in 3D medical images: Application to multiple sclerosis", *Medical Image Analysis*, vol. 6, no. 2, pp. 163–179 *(Cited on page 8)*

Rich, B. G., 2007, "Physical security high-tech niche: outdoor perimeter security sensor technology", retrieved from www.asmag.com/asm/common/ *(Cited on pages 1 and 2)*

Roux, M., 1996, "Automatic registration of SPOT images and digitized maps", *Proceedings of the IEEE International Conference on Image Processing*, Lausanne, Switzerland, pp. 625-628 *(Cited on page 45)*

Russ, John C., 2002, "The Image Processing Handbook", Fourth Edition, CRC Press LCC, Florida *(Cited on pages 4, 8, 22, 63 and 73)*

Sato, S., Arai, Y., Hirota, K., 2000, "Pattern recognition using fuzzy inference with lacked input data", *The Ninth IEEE International Conference on Fuzzy Systems*, vol. 1, no. 7-10, pp. 100-104 *(Cited on page 97)*

Scharstein, D., and Szeliski, R., 2002, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms", *International Journal of Computer Vision*, vol. 47, no. 1-3, pp. 7-42 *(Cited on page 73)*

Senstar-Stellar and Magal Security Systems, 2007, "New paradigm in perimeter security", retrieved from www.asmag.com/asm/common/ *(Cited on page 1)*

Sester, M., Hild, H., and Fritsch, D., 1998, “Definition of ground control features for image registration using GIS data”, *Proceedings of the Symposium on Object Recognition and Scene Classification from Multispectral and Multisensor Pixels*, Columbus, Ohio, pp. 1-7 *(Cited on page 45)*

Shen, J., 2004, “Motion detection in color image sequence and shadow elimination”, in *Visual Communications and Image Processing*, 5308, pp. 731–740, SPIE, (San Jose, USA) *(Cited on page 19)*

Skifstad, K., and Jain, R., 1989, “Illumination independent change detection for real world image sequences”, *Comput. Vis. Graph. Image Process.*, vol. 46, no. 3, pp. 387-399 *(Cited on page 18)*

Stauffer, C., and Grimson, W. E. L., 2000, “Learning patterns of activity using real-time tracking”, *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, no. 8, pp. 747–757 *(Cited on page 8)*

Su, M.S., Hwang, W.L., and Cheng, K.Y., 2004, “Analysis on multiresolution mosaic images”, *IEEE Transactions on Image Processing*, vol. 13, no. 7, pp. 952-959 *(Cited on page 9)*

Tanjung, G., and Lu, T.-F., 2007, “A study on indoor automatic change detection for a mobile-camera”, *Australasian Conference on Robotics and Automation 2007 (ACRA2007)*, Brisbane, Australia *(Cited on pages 44)*

Tanjung, G., Lu, T.-F., Lozo, P., and Liddy, T., 2008, ”A robust approach for detecting the edges of outdoor wire fences”, *2008 Australasian Conference in Robotics and Automation (ACRA 2008)*, Canberra, Australia *(Cited on pages 117 and 129)*

Venkateswar, V., and Chellappa, R., 1995, “Hierarchical Stereo and Motion Correspondence Using Feature Groupings”, *Int’l J. Computer Vision*, vol. 15, pp. 245-269 *(Cited on pages 13 and 23)*

Wang, W., Bergholm, F., and Yang, B., 2003, “Froth delineation based on image classification”, *Mineral Engineering*, vol. 16, Issue 11, pp. 1183-1192 *(Cited on pages 23, 114 and 115)*

Wren, C. R., Azarbayejani, A., Darrell, T., and Pentland, A., 1997, “Pfinder: Real-time tracking of the human body”, *IEEE Trans.Pattern Anal. Machine Intell.*, vol. 19, no. 7, pp. 780–785 *(Cited on page 8)*

Zitnick, C. L., and Kanade, T., 2000, "A Cooperative Algorithm for Stereo Matching and Occlusion Detection”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 7, pp. 675-684 *(Cited on pages 6, 15, 23, 74, 75 and 92)*

Zitova, B., and Flusser, J. 2003, “Image registration methods: a survey”, *Image and Vision Computing*, pp. 977-1000 *(Cited on pages 9 and 22)*