

3D Reconstruction from a Single Still Image Based on Monocular Vision of an Uncalibrated Camera

Tao Yu¹, Jian-Hua Zou², Qin-Bao Song³

^{1,2}Systems Engineering Institute, School of Electronic & Information Engineering, Xi'an Jiaotong University, 28 Xian-Ning West Road, Xi'an city, Shaanxi Province, 710049, P R China

^{1,2}State key laboratory for systems engineering, Xi'an Jiaotong University, 28 Xian-Ning West Road, Xi'an city, Shaanxi Province, 710049, P R China

^{1,3}Dept. of Computer Science & Technology, School of Electronics & Information Engineering, Xi'an Jiaotong University, 28 Xian-Ning West Road Xi'an city, Shaanxi Province, 710049, P R China

¹yvt9399@stu.xjtu.edu.cn, ²jhzou@sei.xjtu.edu.cn, ³qbsong@mail.xjtu.edu.cn

Abstract: we propose a framework of combining Machine Learning with Dynamic Optimization for reconstructing scene in 3D automatically from a single still image of unstructured outdoor environment based on monocular vision of an uncalibrated camera. After segmenting image first time, a kind of searching tree strategy based on Bayes rule is used to identify the hierarchy of all areas on occlusion. After superpixel segmenting image second time, the AdaBoost algorithm is applied in the integration detection to the depth of lighting, texture and material. Finally, all the factors above are optimized with constrained conditions, acquiring the whole depthmap of an image. Integrate the source image with its depthmap in point-cloud or bilinear interpolation styles, realizing 3D reconstruction. Experiment in comparisons with typical methods in associated database demonstrates our method improves the reasonability of estimation to the overall 3D architecture of image's scene to a certain extent. And it does not need any manual assist and any camera model information.

1. Introduction

We all know that the 3D scene has more valuable information than 2D scene because of including the depth information. However, the acquisition for 3D scene is usually complex and difficult. It often needs many cameras to take pictures at multi-view angles and needs the techniques of camera's calibration, image matching etc or needs the complex sophisticated depth sensor instruments to realize.

In reality, sometimes, we cannot acquire many images from different view angles to reconstruct 3D scenes because of the limited conditions, such as the vehicles in the distance, ships in the sea, or the stars in the interstellar space etc, but we are still interested in understanding the hierarchy relationships of front and back among the objects. At these situations, we only can hope to utilize the information from monocular vision to recover or approximate the real 3D effects directly.

Recently, many investigations attempting to reconstruct 3D scenes from monocular vision directly involve mainly two types: from videos [1-5] and from a still image. Among the latter: some demand camera calibration in advance[6-8], some need manual assistance[9-11], some recur to auxiliary accurate

sensors[12], some use the means of camera's focus[13, 14], some append a few extra constrained conditions[15-17].

At present, with the rapid development of Machine Learning, some investigators utilize these theories for training associated models such as Bayes[18, 19], Adaboost[20-23] and Markov random field(MRF)[24-26] etc to estimate the structural information of scenes. However, pure machine learning often brings the problem of under-fitting or over-fitting, which often leads to incorrect or inaccurate estimation for the whole 3D structure of scenes. To make up the deficiency, since optimization's theory and methods can deal with selecting the best alternative in the sense of the given objective function, we consider of introducing the strategy of optimization upon the Machine Learning for estimating the final whole 3D structure of scenes.

In this paper, a method of combining Machine Learning with Dynamic Optimization is used to reconstruct 3D scene directly from a single still image based on monocular vision of an uncalibrated camera. It improves the reasonability of estimation to the whole 3D structure of images' scenes to a certain extent and does not need any manual assist and any camera model information.

The remainder of this paper is arranged as follows. Section 2 interprets the principle of 3D reconstruction from a single still image based on monocular vision of

an uncalibrated camera; Section 3 provides relative experiment and results analysis; Section 4 concludes this paper.

2. The Principle of 3D Reconstruction From A Single Still Image Based on Monocular Vision of An Uncalibrated Camera

2.1 Global Principle

Comparing with 3D image, the monocular vision produces 2D image, which has only the information at orientations of X and Y, without Z. That is to say, it has not the depth information. So, in order to realize 3D reconstruction, we need to estimate the depth information from 2D image.

If we close our one eye, with the other one open to see a picture, we still can feel the relationship of different parts' back and forth to the object in image. To the complex image, according to some rules, we often can also infer its relationship of back or forth with other objects around and further infer the whole architecture of depth from an image. These phenomena provide us with some possibilities for the 3D reconstruction of the scene from monocular image. Here, we think that the perception above maybe some past experience and the rules above maybe some kind of optimization. So, this paper will attempt to combine Machine Learning with Dynamic Optimization for the computer to solve this problem automatically.

In a picture, usually we will see some familiar or unfamiliar object, which involves shape, material, texture, color, and the effect of illumination. And we usually can also accept the inferences that:

A. For a familiar object, if it only can be seen part shape, it is usually sheltered by something.

B. For the material, the part near to us is usually clearer and rougher than the part in distance.

C. For the texture, usually, the near part presents sparser, however, the part in distance presents denser.

D. For the effect of lighting, the part near light source is usually brighter than the part far away.

E. For the color, it usually changes with different objects or different parts of one object or even different lighting sources etc.

According to the experiences above: we can use some samples owning different depth levels for the material, texture and brightness of pictures respectively to train the relative learning machines; then we use some priority to infer the depth level of some area in image and for some objects with familiar shapes, we can use some decision algorithm to infer the possible existed occlusions; finally, to the whole image, we can use optimization algorithm to integrate with all the inferences above; thus, we will acquire multi-possible architectures of depths for an image and we will select the most possible one as the final result according to the computational optimal value or the biggest probability value. In addition, because we do not use any camera's

model information, here, the depths studied are not real absolute depths, but the relative depths of the different components in an image.

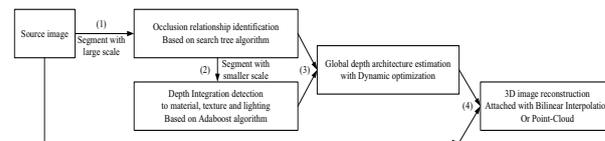


Fig.1 the principle of the 3D reconstruction

Fig.1 gives this paper' main principle of the 3D reconstruction directly from monocular 2D image, in which: the source image is segmented with bigger scale and smaller scale in turn, to make occlusion identification and depth integration detection on material, texture and lighting; then the global depth architecture is estimated with dynamic optimization based on the two stages' results above; at last, combining the global depth architecture with source image, the 3D image reconstruction attached with bilinear interpolation or point-cloud is realized. Next, we will describe each part of the principle in detail.

2.2 Implementation

To estimate the depth structure reasonably, the key is need to make correct image representation as much as possible to the image's important information. Image representation is usually hierarchical, so it is carried on the basis of the image segmentation in different scales [27, 28], in which the image's different components represent different meanings.

2.2.1 Occlusion Identification

First, we segment the image in bigger scale with method in Ref. [27] to acquire familiar objects and identify the associated occlusions. Here, the occluding phenomenon is defined as: when object A can only be seen in partial shape and the other area where should be embodied as the rest shape of A, is occupied by object B, then we think that object A is occluded by object B, or B is in front of A. The occluded areas' divisions refer to [29]. The related principle is as Eqs. (2-1) displaying, in which M, N are two neighboring objects in an image; X represents the associated traits; Y is the associated prior knowledge for familiar objects and Q is the number of objects in an image. According to Bayesian rule, the more remarkable the traits information $P(X/(M, N))$ and the prior information $P(Y/(M, N, X))$ are, the bigger the value of posterior probability function $P((M, N)/(X, Y))$ on occlusion will be.

$$P((M, N)/(X, Y)) = \frac{P(Y/(M, N, X))P(X/(M, N))}{\sum_{M \neq N \in Q} [P(Y/(M, N, X))P(X/(M, N))]} \quad (2-1)$$

To find out all the familiar objects' occlusive relationships in an image, we adopt a kind of searching tree algorithm. The searching process is just as Fig.2, in which all the familiar objects' relative occlusive relationships of each other are inferred by the principle

of Eqs. (2-1) and at last, relative different number is used to represent different depth hierarchy. Of course, there exists the likelihood of inferring error on the occluding problem. It relies on whether the prior information on the familiar objects' understanding is reliable. And the more reliable the prior knowledge is, the more accurate the inference is.

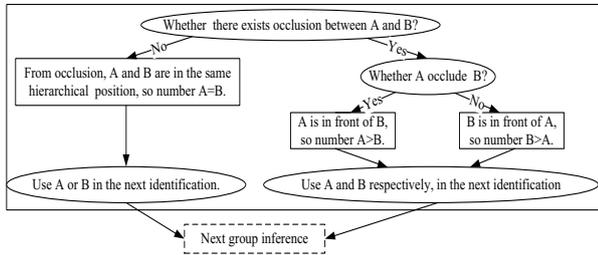


Fig.2 the search strategy on occlusion in one image

2.2.2 Depth Integration Detection to Material, Texture and Lighting

After the occlusion identification is over, then, the image will be segmented again in smaller scale with method in Ref. [27] to acquire the superpixels, each of which represents a coherent region with similar properties such as lighting, texture, material or colors. Fig.3 displays an example of segmenting image, in which there are two grades of segmentation with yellow color and red color describing respectively.



3-a Source image 3-b Superpixel segmenting image

Fig.3 An example of segmenting image

Next, we process the superpixels' image above from the aspects of lighting, texture and material to extract more depth information. Here, before further depth identification, we predefine multiple depth classes according to different changing extent of lighting, texture and material with depth changing. We use the logistic regression version of AdaBoost algorithm in Ref. [30] to train each class of depth samples and acquire relative detector. At each class of depth, the relative detector based on textural or material traits' weighting is trained and the relative lighting model is also trained. In identifying, according to some priority order, we integrate the detectors to acquire the integration depth on material, texture and lighting. Fig.4 displays the associated detecting principle, in which the property of AdaBoost that uses weak classifiers to construct strong classifiers is applied for both the single depth class detection and the integration depth detection.

On the selection of associated traits: the textural energy which are computed from Law's masks in Refs. [31, 32] act as textural features to embody the extent of textural denseness at different depths; the Haar-like traits in Ref. [33] act as material features to describe the extent of material smoothness and ambiguity with depth changing; the lighting model on depth refers to [34, 35], in which the distance from light source to viewer is regarded as the relative depth herein, then image's other components are assigned to corresponding lighting categorical depth classes and model parameters in turn.

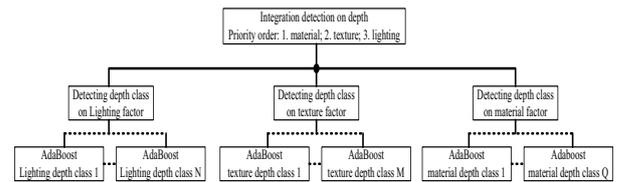


Fig.4. The principle of integration detection to the depth of lighting, texture and material

2.2.3 Global Depth Architecture

To acquire the final global depth architecture of an image, we combine with all the inferences of the factors above and apply a kind of dynamic optimization with constrain conditions to solve the problem. In addition, image's gray values usually reflect some depth information, so it will also be regarded as one base applied in the estimation for the final global depth architecture. The corresponding formula is Eqs. (2-2), where: m, n represents two neighboring areas in an image; $g(m,n,d)$ is the depth mapping function between neighboring area m and n corresponding to the constrains above. Herein: $q=1$, base on gray image; $q=2$, base on lighting; $q=3$, base on texture; $q=4$, base on material; $q=5$, base on occluding; λ and μ are associated Lagrange multipliers; L is the total number of correlations corresponding to neighboring areas in an image; d is the final depth correlation to be established for all the smaller areas of an image; c is introduced slack variable used to transform the inequality constrains above to equality constrains.

$$\arg \min_{d, c} \left(\sum_{p=1}^L d_p - \sum_{q=1}^5 (\lambda_{pq} \cdot (g_q(m_p, n_p, d_p) - c_{pq})) + \frac{1}{2\mu_p} \cdot \sum_{q=1}^5 ((g_q(m_p, n_p, d_p) - c_{pq})^2) \right) \quad (2-2)$$

From optimization theory in Ref. [36], we can infer optimal

the value of c_{pq} as Eqs. (2-3)

$$c_{pq} = \max(g_q(m_p, n_p, d_p) - \mu_p \cdot \lambda_{pq}, 0) \quad (2-3)$$

Combing Eqs. (2-2) with (2-3), we can construct the augmented Lagrangian function (2-4).

$$LA(d, \lambda, \mu) = \sum_{p=1}^L \left(d_p - \sum_{q=1}^5 (\lambda_{pq} \cdot (g_q(m_p, n_p, d_p) - \max(g_q(m_p, n_p, d_p) - \mu_p \cdot \lambda_{pq}, 0))) + \frac{1}{2\mu_p} \cdot \sum_{q=1}^5 ((g_q(m_p, n_p, d_p) - \max(g_q(m_p, n_p, d_p) - \mu_p \cdot \lambda_{pq}, 0))^2) \right) \quad (2-4)$$

Thus the problem (2-2) above is transformed into formula (2-5).

$$\min_d (LA(d, \lambda, \mu)) \quad (2-5)$$

Referring to [37], the specific solving procedures on problem

the above are as follow:

Given $\mu_0 > 0$, tolerance $\varepsilon > 0$, starting points \mathbf{d}^{s0} and λ^0 ;

for $k = 0, 1, 2, \dots$

Find an approximate minimizer \mathbf{d}^k of $LA(\bullet, \lambda^k; \mu^k)$, starting at \mathbf{d}^{sk} ,

and terminating when $\|\nabla_d LA(\mathbf{d}^k, \lambda^k; \mu^k)\| \leq \varepsilon$;

if final convergence test for formula (2-5) is satisfied

Stop with approximate solution \mathbf{d}^k ;

end(if)

Update Lagrange multipliers using Eqs. (2-6) to obtain λ^{k+1} ;

$$\lambda^{k+1} = \max(\lambda^k - g(m, n, d^k) / \mu^k, 0); \quad (2-6)$$

Choose new penalty parameter $\mu^{k+1} \geq \mu^k$;

Set starting point for the next iteration to $\mathbf{d}^{s(k+1)} = \mathbf{d}^k$;

end (for)

According to the method described above, the vector \mathbf{d} , λ and μ are integrated updating until convergence. Note that the vector \mathbf{d} refers to relative depth, not absolute depth.

Fig.5 illustrates an example of the estimation to an image's global depth architecture with the method of this paper, in which the different gray levels indicate different depths.

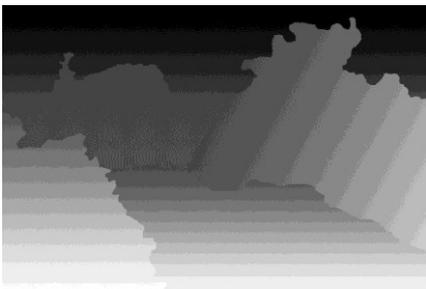
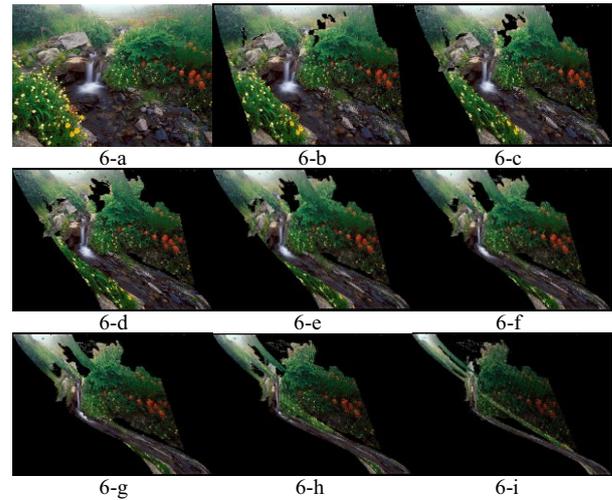


Fig.5 An example of the estimation to the global depth architecture of an image

2.2.4 3D Reconstruction Realization

At last, after acquiring the complete depthmap of an image, we combine the depthmap with its source image to reconstruct the 3D image. And here, two kinds of methods are used to process the voxels in 3D image: to the images distributed mainly with continuous qualities, the bilinear interpolation method is applied to complementing and smoothing the color area among the reconstructed 3D pixel points; to the images distributed mainly with discrete qualities, the point-cloud method is applied directly to the reconstruction of the 3D image's voxels. In addition, during the course of 3D image reconstruction, to the borderlines of the segmentation before, we use the corresponding pixels in original image to replace them. Fig.6 gives an example of reconstructed 3D images viewed in different visual angles from a single 2D source image. From different

view angles, we can see that they all approximate the usual real objects and scenes.



a-frontal image; b- right rotation 10° ; c- right rotation 20° ; d- right rotation 30° ; e- right rotation 40° ; f- right rotation 50° ; g- right rotation 60° ; h- right rotation 70° ; i- right rotation 80° ;

Fig.6 Example of 3D reconstruction from single image

3. Experiment

3.1 Testing Environment

In order to prove the ability of the application in general environment for the method proposed in this paper, all the experiments are conducted in the environment of Microsoft visual c++ 6.0 at the platform of Pentium 1.73GHz personal computer.

The framework of 3D reconstruction from monocular vision proposed by this paper is tested with the Make3D Range Image Data' dataset I^[24, 25], which consists of 534 image+depthmap pairs, with an image resolution of 2272×1704 and a depthmap resolution of 55×305. In our experiment, 400 of the images/depthmaps are used for training associated detectors and the remaining 134 for testing. At the same time, we will compare the experimental results with typical Saxena^[24] method and HEH^[23] method. And for fairness, their depthmaps are scaled and shifted before computing the errors to match the global scale of our test images.

3.2 Testing Results and Analysis

Fig.7 gives some examples of real effects for contrasting typical methods —Saxena^[24] and HEH^[23] with ours. In comparisons: HEH's generating a popup effect by folding the images at "ground-vertical" boundaries is not feasible for all the images. Especially for multy non-ground and discontinuous areas with front and back relationships in an image, it often fails to

describe these scenes' structure correctly, such as fig.7-b1, 7-b2. Saxena's using Markov Random Field (MRF) to infer both 3D location and orientation of the patches in an image without any explicit assumptions about the structure of the scene makes it generalize well, even to scenes with significant nonvertical structure. The deficiency exists in that at the aspect of the whole structural reconstruction, it lacks the comprehensive considering for many factors, so its global 3D structural estimation is sometimes inaccurate, such as fig. 7-c3. Our algorithm's global dynamic optimization seems to be more reasonable and feasible for the whole 3D architectural reconstruction to a certain extent, producing better effects although a few small local parts in images are estimated somewhat coarsely.



a-Source image; b-HEH' algorithm; c-Saxena's algorithm; d-our algorithm

Fig.7 Examples of 3D reconstruction from single image with different algorithms

Table.1 and Fig.8 give the final statistical results for the testing comparisons of our algorithm with the two typical algorithms above, in which: whether 'Depth error' or 'Relative Depth error' are both averaged over all pixels in the hold-out test set; 'Correct rate' means percent of models qualitatively correct; 'Planes Correct rate' is percent of major planes correctly identified. Combing with the analysis before, we can see that from the whole testing data, Saxena's method is a little better than ours probably because of our method's accumulated several local small estimation errors during the course of computation and our method outperforms HEH mainly for HEH' sometimes failed estimation for the whole structure of scenes, all of which verifies our algorithm's feasibility and rationality further.

Table.1 comparisons with typical methods

Method	Depth error $ \log_{10}d - \log_{10}\hat{d} $	Relative Depth error $ \hat{d} - d /d$	Correct rate (%)	Planes Correct rate (%)
HEH	0.320	0.423	33.1	50.3
Saxena	0.187	0.370	64.9	71.2
Ours	0.253	0.382	52.5	69.4

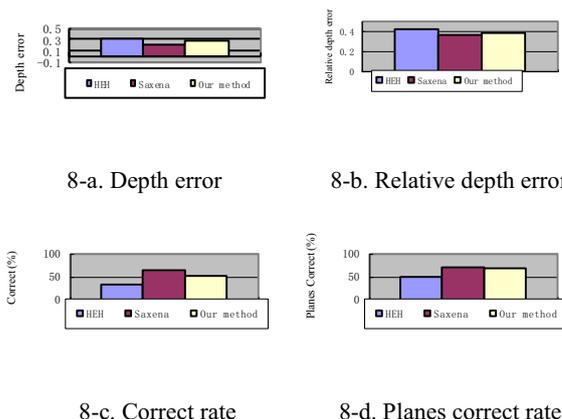


Fig.8 comparisons with typical methods.

4. Conclusion

This paper mainly proposes a framework of 3D reconstruction from a single still image of unstructured outdoor environment based on monocular vision of an uncalibrated camera. It integrates machine's learning for each factor with dynamic optimization's considering comprehensively for multy-factors to estimate scene's 3D structure and does not need any manual assist, any camera model information. In experiment, we use Make3D Range Image Data to compare it with typical methods of HEH and Saxena. The results basically confirm our method's reasonability and feasibility for image scene's global architectural 3D reconstruction to a certain extent.

In the next work, our investigation will manage to improve the accuracy of estimation not only for the scene's local patches especially, but also for the global structure further. Meanwhile, whether this paper's method can apply to other databases needs new testing. In addition, we will investigate the real-time 3D reconstruction from a single still image, which needs paralleled mechanism and improving our algorithm's speed furtherly.

Acknowledgment

The authors thank to Computer Science Department, Stanford University for supplying investigators with the free Make3D Range Image Data and thank for being supported by the National Natural Science Foundation of China (50177025).

References

1. Jebara, T., A. Azarbayejani, and A. Pentland, 3D structure from 2D motion. Signal Processing magazine, 1999. **16**(3): p. 66-84.
2. Po, L.-M., et al., Automatic 2D-to-3D video conversion technique based on depth-from-motion and color segmentation, in Signal Processing (ICSP), 2010 IEEE 10th International Conference

- on. 2010, IEEE. p. 1000-1003.
3. Bok, Y., Y. Hwang, and I.S. Kweon, Accurate motion estimation and high-precision 3d reconstruction by sensor fusion, in *Robotics and Automation, 2007 IEEE International Conference on*. 2007, IEEE. p. 4721-4726.
 4. Cao, X., Z. Li, and Q. Dai, Semi-automatic 2D-to-3D conversion using disparity propagation. *IEEE TRANSACTIONS ON BROADCASTING*, 2011. **57**(2): p. 491-499.
 5. Hertzmann, A. and S.M. Seitz, Example-based photometric stereo: Shape reconstruction with general, varying brdfs. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 2005. **27**(8): p. 1254-1264.
 6. Guillou, E., et al., Using vanishing points for camera calibration and coarse 3D reconstruction from a single image. *The Visual Computer*, 2000. **16**(7): p. 396-410.
 7. Wilczkowiak, M., E. Boyer, and P. Sturm. Camera calibration and 3D reconstruction from single images using parallelepipeds. in *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*. 2001. IEEE.
 8. Wang, G., et al., Camera calibration and 3D reconstruction from a single view based on scene constraints. *Image and Vision Computing*, 2005. **23**(3): p. 311-323.
 9. Criminisi, A., I. Reid, and A. Zisserman, Single View Metrology. *International Journal of Computer Vision*, 2000. **40**(2): p. 123-148.
 10. Barinova, O., et al., Fast automatic single-view 3-d reconstruction of urban scenes, in *Computer Vision-ECCV 2008*. 2008, Springer. p. 100-113.
 11. Sturm, P. and S. Maybank. A method for interactive 3d reconstruction of piecewise planar objects from single images. in *The 10th British Machine Vision Conference (BMVC'99)*. 1999. The British Machine Vision Association (BMVA).
 12. Willneff, J., J. Poon, and C. Fraser, Single-image high-resolution satellite data for 3D information extraction. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2005. **36**(1/W3): p. 1-6.
 13. Namboodiri, V.P. and S. Chaudhuri, Recovery of relative depth from a single observation using an uncalibrated (real-aperture) camera, in *Computer Vision and Pattern Recognition, CVPR. IEEE Conference on*. 2008, IEEE. p. 1-6.
 14. Zhuo, S. and T. Sim, On the recovery of depth from a single defocused image, in *Computer Analysis of Images and Patterns*. 2009, Springer. p. 889-897.
 15. Van den Heuvel, F.A., 3D reconstruction from a single image using geometric constraints. *ISPRS Journal of Photogrammetry and Remote Sensing*, 1998. **53**(6): p. 354-368.
 16. El-Hakim, S.F. A flexible approach to 3D reconstruction from single images. in *ACM SIGGRAPH*. 2001. Citeseer.
 17. Wang, G., et al., Single view metrology from scene constraints. *Image and Vision Computing*, 2005. **23**(9): p. 831-840.
 18. Delage, E., H. Lee, and A.Y. Ng, A dynamic bayesian network model for autonomous 3d reconstruction from a single indoor image, in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*. 2006, IEEE. p. 2418-2428.
 19. Han, F. and S.-C. Zhu. Bayesian reconstruction of 3d shapes and scenes from a single image. in *Higher-Level Knowledge in 3D Modeling and Motion Analysis, 2003. HLK 2003. First IEEE International Workshop on*. 2003. IEEE.
 20. Hoiem, D., A.A. Efros, and M. Hebert, Recovering surface layout from an image. *International Journal of Computer Vision*, 2007. **75**(1): p. 151-172.
 21. Hoiem, D., A.A. Efros, and M. Hebert, Closing the loop in scene interpretation, in *Computer Vision and Pattern Recognition. CVPR. IEEE Conference on*. 2008, IEEE. p. 1-8.
 22. Hoiem, D., A.A. Efros, and M. Hebert, Automatic photo pop-up. *ACM Transactions on Graphics (TOG)*, 2005. **24**(3): p. 577-584.
 23. Hoiem, D., A.A. Efros, and M. Hebert, Geometric context from a single image, in *Computer Vision. ICCV 2005. Tenth IEEE International Conference on*. 2005, IEEE. p. 654-661.
 24. Saxena, A., M. Sun, and A.Y. Ng, Make3d: Learning 3d scene structure from a single still image. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 2009. **31**(5): p. 824-840.
 25. Saxena, A., S.H. Chung, and A.Y. Ng, Learning depth from single monocular images, in *Advances in Neural Information Processing Systems(NIPS)*. 2005. p. 1161-1168.
 26. Delage, E., H. Lee, and A.Y. Ng, Automatic single-image 3d reconstructions of indoor manhattan world scenes, in *Robotics Research*. 2007, Springer. p. 305-321.
 27. Levinshtein, A., et al., Turbopixels: Fast superpixels using geometric flows. 2009, *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*,. p. 2290-2297.
 28. Felzenszwalb, P.F. and D.P. Huttenlocher, Efficient graph-based image segmentation, in *International Journal of Computer Vision*. 2004, Springer. p. 167-181.
 29. Hoiem, D., et al., Recovering occlusion boundaries from a single image, in *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*. 2007, IEEE. p. 1-8.
 30. Collins, M., R.E. Schapire, and Y. Singer, Logistic regression, AdaBoost and Bregman distances, in *Machine Learning*. 2002, Springer. p. 253-285.
 31. Davies, E.R., Laws' texture energy in texture. 1997, *Machine vision: theory, algorithms, practicalities*.

32. Laws, K.I., Rapid texture identification, in 24th Annual Technical Symposium. 1980, International Society for Optics and Photonics. p. 376-381.
33. Lienhart, R., A. Kuranov, and V. Pisarevsky, Empirical analysis of detection cascades of boosted classifiers for rapid object detection, in Pattern Recognition. 2003, Springer. p. 297-304.
34. Klassen, R.V., Modeling the effect of the atmosphere on light. 1987, ACM. p. 215-237.
35. Narasimhan, S.G. and S.K. Nayar, Shedding light on the weather, in Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on. 2003, IEEE. p. I-665-I-672 vol. 1.
36. Nocedal, J. and S.J. Wright, Numerical Optimization. Springer, 2006. **9**(4): p. 1556-1556.
37. Nocedal, J. and S.J. Wright, Numerical Optimization Second Edition. 1999..