

STUDY ON THE FEASIBILITY OF RGB SUBSTITUTE CIR FOR AUTOMATIC REMOVAL VEGETATION OCCLUSION BASED ON GROUND CLOSE-RANGE BUILDING IMAGES

Chang LI^{a*}, Fangfang LI^b, Yawen LIU^c, Xi LI^d, Pengcheng LIU^a, Benlin XIAO^e

^a College of Urban and Environmental Science, Central China Normal University, Wuhan 430079, China
lcshaka@126.com && lichang@mail.ccnu.edu.cn

^b Science and Technology on Information Systems Engineering Laboratory, National University of Defense Technology, Changsha 410073, China

^c School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China

^d State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China.

^e Civil Engineering & Architecture School, Hubei University of technology, Wuhan 430068, China

Commission V, V/6

KEY WORDS: CIR, RGB, Vegetation occlusion, Removal, Segmentation, 3D reconstruction

ABSTRACT:

Building 3D reconstruction based on ground remote sensing data (image, video and lidar) inevitably faces the problem that buildings are always occluded by vegetation, so how to automatically remove and repair vegetation occlusion is a very important preprocessing work for image understanding, compute vision and digital photogrammetry. In the traditional multispectral remote sensing which is achieved by aeronautics and space platforms, the Red and Near-infrared (NIR) bands, such as NDVI (Normalized Difference Vegetation Index), are useful to distinguish vegetation and clouds, amongst other targets. However, especially in the ground platform, CIR (Color Infra Red) is little utilized by compute vision and digital photogrammetry which usually only take true color RGB into account. Therefore whether CIR is necessary for vegetation segmentation or not has significance in that most of close-range cameras don't contain such NIR band. Moreover, the CIE L^*a^*b color space, which transform from RGB, seems not of much interest by photogrammetrists despite its powerfulness in image classification and analysis. So, CIE (L, a, b) feature and support vector machine (SVM) is suggested for vegetation segmentation to substitute for CIR. Finally, experimental results of visual effect and automation are given. The conclusion is that it's feasible to remove and segment vegetation occlusion without NIR band. This work should pave the way for texture reconstruction and repair for future 3D reconstruction.

1. INTRODUCTION

Photo-realistic 3D models are nowadays required in many applications (Ortin and Remondino, 2005). Building 3D-reconstruction is an important part of work in Digital City and also is a significant part of street landscape, therefore how to carry out 3D visualization plays an important role in 3D city modelling (LI and Zhou, 2010). Terrestrial images for texture mapping streetscapes, which are captured along a narrow street, are hardly free of the occlusions which hinder the realistic façade texture. Thus, the full automation of texture occlusion removal has been regarded as a valuable research for rendering street scene (LIU and GUAN, 2010). Occlusions can be moving (e.g. pedestrians) or static (e.g. monuments) objects, which occlude the full and free visibility of the surface to be textured (Ortin and Remondino, 2005). This paper mainly discusses the static problem that building is occluded by vegetation.

In the traditional multispectral remote sensing which is achieved by aeronautics and space platforms, the Red and Near-infrared (NIR) bands, such as NDVI (Normalized Difference Vegetation Index), are useful to distinguish vegetation and clouds, amongst other targets. However, especially in the ground platform, NIR band is little utilized by compute vision

and digital photogrammetry which usually only take visible light RGB bands into account. Therefore, in the ground close-range scene of buildings, whether NIR is necessary to recognize vegetation occlusion is discussed and tested.

2. VEGETATION OCCLUSION SEGMENT AND REMOVAL

2.1 Vegetation segment and removal of CIR image

Color infrared image (CIR) is not as popular as black and white for pictorial use, but it does have its place. Color infrared image is sometimes referred to as false color. This is because this image is designed to differentiate between colors rather than reproduce them accurately. And CIR is very important for data acquisition and updating, especially for vegetation (LYON et al., 1998).

The Normalized Difference Vegetation Index (NDVI) is a simple graphical indicator that can be used to analyze remote sensing measurements, typically but not necessarily from a space platform, and assess whether the target being observed contains live green vegetation or not. NDVI was one of the most successful of many attempts to simply and quickly identify

* Corresponding author.

vegetated areas and their "condition," and it remains the most well-known and used index to detect live green plant canopies in multispectral remote sensing data. Once the feasibility to detect vegetation had been demonstrated, users tended to also use the NDVI to quantify the photosynthetic capacity of plant canopies. The NDVI is calculated from these individual measurements as follows (NASA) :

$$NDVI = \frac{NIR - VIS}{NIR + VIS} \quad (1)$$

These spectral reflectances are themselves ratios of the reflected over the incoming radiation in each spectral band individually; hence they take on values between 0.0 and 1.0. By design, the NDVI itself thus varies between -1.0 and +1.0. It should be noted that NDVI is functionally, but not linearly, equivalent to the simple infrared/red ratio (NIR/VIS).

There are many methods that can be used for image segmentation. The NDVI is one of the most widely used indices for differentiating between vegetation and non-vegetation areas in remote sensing (ZHANG et al., 2006). For the NDVI, the threshold for vegetation extraction is usually positive and near to zero, it may vary from 0.05 to 0.15. Human supervision is helpful for selecting the best threshold from typical imagery. For our test data, non-vegetation areas are not well removed with a threshold of 0.0, while many areas are falsely removed with a threshold of 0.2. The best result is obtained with a threshold of 0.1.

2.2 Vegetation segment and removal of RGB image

In the traditional multispectral remote sensing which is achieved by aeronautics and space platforms, the Red and Near-infrared (NIR) bands. However, especially in the ground platform, NIR band is little utilized by compute vision and digital photogrammetry which usually only take RGB bands into account. Therefore, it is important to recognize vegetation occlusion in the ground close-range scene of buildings based on visible light RGB images.

CIE L*a*b is very popular for image segmentation in computer vision. This is based directly on CIE XYZ (1931) and is another attempt to linearise the perceptibility of unit vector colour differences. Again, it is non-linear, and the conversions are still reversible. Colouring information is referred to the colour of the white point of the system, subscript n. The non-linear relationships for L* a* and b* are the same as for CIELUV and are intended to mimic the logarithmic response of the eye (Ford and Roberts, 1998).

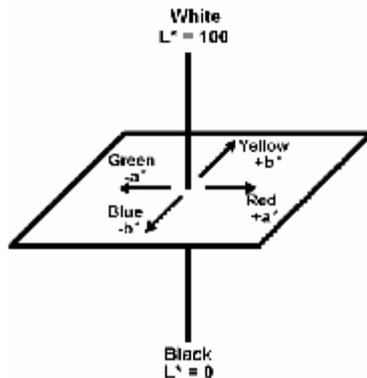


Figure 1. Definition of CIE L*a*b

The CIELAB color scale is an approximately uniform color scale. In a uniform color scale, the differences between points plotted in the color space correspond to visual differences

between the colors plotted. The CIELAB color space is organized in a cube form. The L* axis runs from top to bottom. The maximum for L* is 100, which represents a perfect reflecting diffuser. The minimum for L* is zero, which represents black. The a* and b* axes have no specific numerical limits. Positive a* is red. Negative a* is green. Positive b* is yellow. Negative b* is blue. Fig1 is a diagram representing the CIELAB color space (HunterLab, 2008).

The CIE XYZ colour space was presented by CIE in 1931 (SEII 2002). Three axes X, Y and Z are orthogonally defined by the basic colours R, G and B (red, green and blue). Generally, only points in the surface X+Y+Z=1 are considered. Each RGB point can be transformed into colour space CIE XYZ as follows (SEII, 2002):

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412291 & 0.357664 & 0.180209 \\ 0.212588 & 0.715329 & 0.072084 \\ 0.019326 & 0.119221 & 0.949102 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2)$$

The component L represents the light Lightness with value from 0 to 100. The components a* and b* represent colour: a* varies from green (with value -120) to red (with value + 120), b* varies from blue (with value -120) to yellow (with value + 120).

Those can be written as:

$$\left. \begin{aligned} L &= 116 \cdot (Y / Y_n)^{1/3} - 16 && \text{if } 0.008856 < (Y / Y_n) \\ L &= 903.3 \cdot (Y / Y_n) && \text{else} \\ a &= 500 \cdot (F(X / X_n) - F(Y / Y_n)) \\ b &= 200 \cdot (F(Y / Y_n) - F(Z / Z_n)) \end{aligned} \right\} \quad (3)$$

Where (X, Y, Z) is the point to be converted, which can be obtained from Eq. (2). Let p represents X/X_n, Y/Y_n and Z/Z_n respectively, then:

$$\left. \begin{aligned} F(p) &= p^{1/3} && \text{if } p > 0.008856 \\ F(p) &= 7.787 \cdot p + 16/116 && \text{else} \end{aligned} \right\} \quad (4)$$

(X_n, Y_n, Z_n) is the tristimulus values for the illuminant (HunterLab, 2008), which is also called white point. Here illuminant can be: X_n=0.312779, Y_n=0.329184, Z_n=0.358037 (ZHANG et al., 2006). Other parameters can be (Ford and Roberts, 1998; HunterLab, 2008):

$$\left. \begin{aligned} \Delta L^* &= L^*_{\text{sample}} - L^*_{\text{standard}} \\ \Delta a^* &= a^*_{\text{sample}} - a^*_{\text{standard}} \\ \Delta b^* &= b^*_{\text{sample}} - b^*_{\text{standard}} \\ \Delta E^* &= \sqrt{\Delta L^{*2} + \Delta a^{*2} + \Delta b^{*2}} \\ \Delta C^* &= C^*_{\text{sample}} - C^*_{\text{standard}} \\ h_{ab} &= \arctan(b^* / a^*) \end{aligned} \right\} \quad (5)$$

Where:

$$C^* = \sqrt{a^{*2} + b^{*2}}$$

± Δ L* means sample is lighter or darker than standard; ± Δ a* means sample is redder or greener than standard; ± Δ b* means sample is yellower or bluer than standard

Many experiments have been carried out to compare the performance of image segmentation between NDVI and CIE L*a*b (ZHANG et al., 2006). So the CIE L*a*b approach is used for vegetation segment and removal. Another reason is that using CIE L*a*b, vegetation can also be extracted from visible light RGB images because the component a* is negative for vegetation in standard visible light RGB imagery and close to -120 for green vegetation. To segment RGB imagery with CIE L*a*b, a* from -0.15 to -0.05 should be applied as a threshold.

The best threshold can again be obtained under human supervision. Meanwhile, it is of course not possible for NDVI to deal with RGB imagery.

In order to get better removal result, a supervised learning method, support vector machine (SVM), is utilized for extracting vegetation. We choose (L, a, b) as characteristic variables. Given a set of training examples, each marked as belonging to one of two categories (vegetation or not), an SVM training algorithm builds a model that assigns new examples into one category or the other. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible (BURGES, 1998). Linear SVM can be used in this paper.

An n -dimensional pattern (object) \mathbf{x} has n coordinates, $\mathbf{x}=(x_1, x_2, \dots, x_n)$, where each x_i is a real number, $x_i \in R$ for $i = 1, 2, \dots, n$. Each pattern \mathbf{x}_j belongs to a class $y_j \in \{-1, +1\}$. Consider a training set T of m patterns together with their classes, $T=\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_m, y_m)\}$. Consider a dot product space S , in which the patterns \mathbf{x} are embedded, $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m \in S$. Any hyperplane in the space S can be written as (BURGES, 1998) :

$$\{\mathbf{x} \in S \mid \mathbf{w} \cdot \mathbf{x} + b = 0\}, \quad \mathbf{w} \in S, b \in R \quad (6)$$

The dot product $\mathbf{w} \cdot \mathbf{x}$ is defined by:

$$\mathbf{w} \cdot \mathbf{x} = \sum_{i=1}^n w_i \cdot x_i \quad (7)$$

A training set of patterns is linearly separable if there exists at least one linear classifier defined by the pair (\mathbf{w}, b) which correctly classifies all training patterns. This linear classifier is represented by the hyperplane $H(\mathbf{w} \cdot \mathbf{x} + b = 0)$ and defines a region for class +1 patterns $(\mathbf{w} \cdot \mathbf{x} + b > 0)$ and another region for class -1 patterns $(\mathbf{w} \cdot \mathbf{x} + b < 0)$.

After training, the classifier is ready to predict the class membership for new patterns, different from those used in training. The class of a pattern \mathbf{x}_k is determined with the equation:

$$\text{class}(x_k) = \begin{cases} +1 & \text{if } \mathbf{w} \cdot \mathbf{x}_k + b > 0 \\ -1 & \text{if } \mathbf{w} \cdot \mathbf{x}_k + b < 0 \end{cases} \quad (8)$$

Therefore, the classification of new patterns depends only on the sign of the expression $\mathbf{w} \cdot \mathbf{x} + b$. After selecting samples, we can recognize two classes problem (vegetation or not) properly.

3. EXPERIMENTS AND RESULTS

The aim of this section is to evaluate the feasibility and effectiveness of the proposed occlusion detection technique. The proposed method was implemented by C++.

3.1 CIR image segmentation and removal

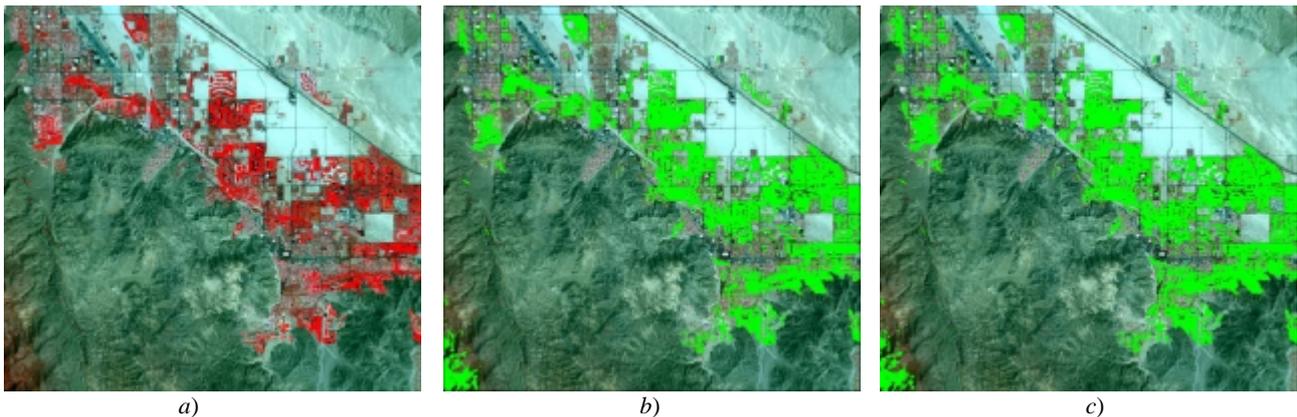


Figure 1. a) Standard false color composite satellite image. b) Result by NDVI. c) Result by CIE L*a*b.

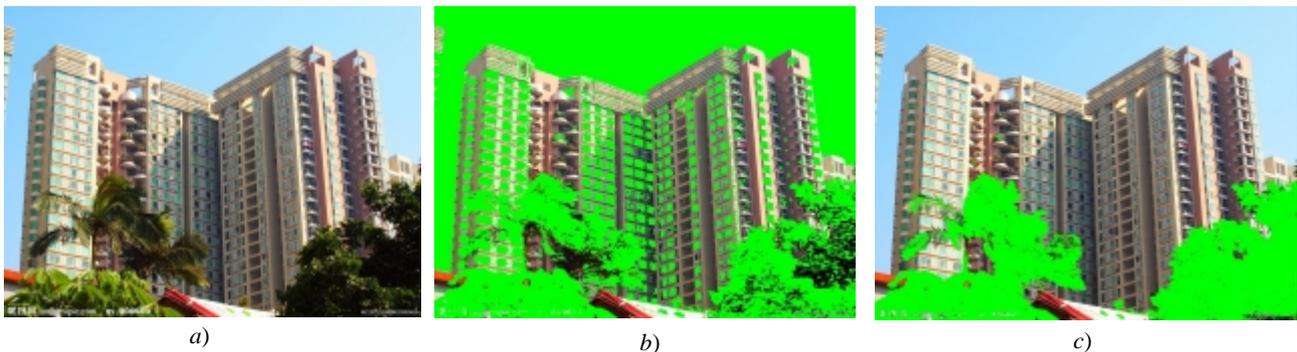


Figure 2. a) True color close-range image. b) Result by CIE L*a*b with a threshold. c) Result by feature (L, a, b) with SVM.

The first experimental data is satellite image that is composed by standard false color. And result of vegetation extraction by NDVI (with a threshold $\text{NDVI} > 0.1$) and CIE L*a*b (with a threshold $a > 12$) are assigned by green color in Fig 1. b) and c). We can find that there are almost the same results between these two methods. So the CIE L*a*b approach is used for vegetation

extraction in this paper. Furthermore CIE L*a*b, vegetation can also be extracted from visible light RGB images because the component a^* is negative, which is tested as follows.

3.2 RGB image segmentation and removal

The second experimental data which is the focus in this study is close-range building RGB image under the ground platform. Because this is visible light RGB image shot by ordinary digital camera, NDVI approach is no longer applicable. CIE L^*a^*b method is used for segment. Under the threshold $a < -12$, removal results is Fig 2 b). The sky and windows are segmented in error. The segment method of satellite image with CIE L^*a^*b

does not apply to ground image, because there are some noise by foreground and background. Vegetation extraction based on SVM is utilized under human supervision, whose result can be seen in Fig 2 c). The green areas in the image are vegetation ones. And the correct recognition rate of vegetation occlusion is not only better than Fig 2 b), but also the wrong recognition rate is reduced greatly.

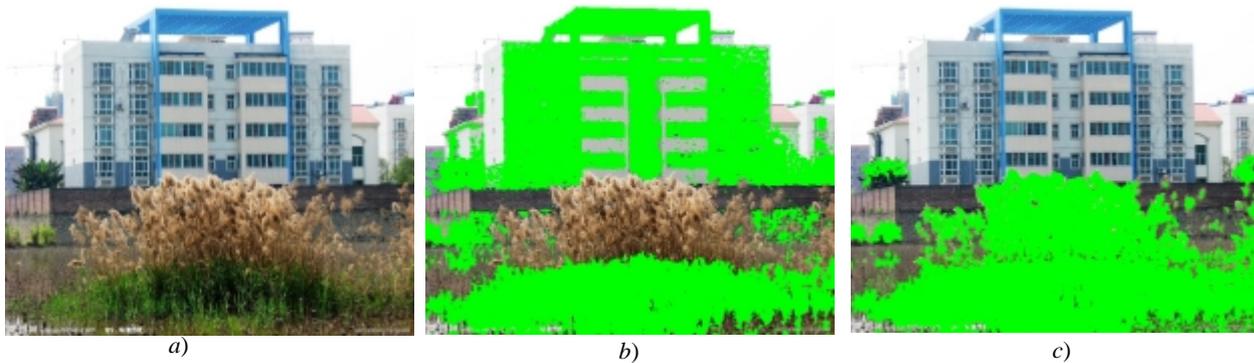


Figure 3. a) True color close-range image. b) Result by CIE L^*a^*b with a threshold. c) Result by feature (L, a, b) with SVM.

The third experimental data is close-range building RGB image where there are yellow or brown vegetations. Under the threshold $a < -12$, there are not only many error segment areas in Fig 3. b), but also some withered and yellow vegetation isn't recognized by this method. Thus, SVM method based on selecting (L, a, b) as feature is implemented for removal vegetation occlusion. The experimental results are shown in Fig 3 c). We can find that almost all the occlusion of vegetation is removed.

3.3 Conclusions and future work

An effective approach for extraction and segment the vegetation occlusion of building by RGB close-range imagery is presented. The proposed image segmentation approach by CIE L^*a^*b and SVM has good potential for the removal of vegetation occlusion from imagery since it works on both CIR and standard RGB imagery. Especially, no matter whether vegetation is green or yellow, it adapts to the segment and removal of ground vegetation occlusion. We can draw the conclusion that RGB image can substitute for CIR image to recognize vegetation occlusion.

Tree that is a familiar occlusion has branches, leaves and trunk, which leads to spectrum difference, texture difference and geometry difference (edge). Hence, in order to get better result of occlusion removal, more information and feature (spectrum, texture and geometry) should be implemented and synthesized. This work should pave the way for texture reconstruction and repair for future 3D reconstruction.

3.4 Acknowledgements and Appendix (optional)

This work is supported by National Natural Science Foundation of China (NSFC) (Grant No. 41101407, 41001204 and 41001260), the Natural Science Foundation of Hubei Province, China (Grant No. 2010CDZ005) and self-determined research funds of CCNU from the colleges' basic research and operation of MOE (Grant No. CCNU10A01001). Heartfelt thanks are also given for the comments and contributions of anonymous reviewers and members of the Editorial team.

References

- [1] BURGESS, C.J.C., 1998. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*(2): 121–167.
- [2] Ford, A. and Roberts, A., 1998. Colour Space Conversions - <http://www.poynton.com/PDFs/coloureq.pdf>.
- [3] HunterLab, 2008. CIE $L^*a^*b^*$ Color Scale - www.hunterlab.com/appnotes/an07_96a.pdf.
- [4] LI, C. and Zhou, Y., 2010. 3D Auto-Reconstruction for Street Elevation Based on Line and Plane Feature, *The 2nd International Conference on Computer and Automation Engineering*, Singapore, pp. 460-466.
- [5] LIU, Y. and GUAN, Z., 2010. A GRID-BASE LINE ANALYSIS FOR STREET OCCLUSION REMOVAL AND BUILDING FACADE TEXTURING, *ASPRS Annual Conference*, San Diego, California.
- [6] LYON, J.G., YUAN, D., LUNETTA, R.S. and ELVIDGE, C.D., 1998. A change detection experiment using vegetation indices. *Photogrammetric Engineering and Remote Sensing*, 66: 143–150.
- [7] NASA, Normalized Difference Vegetation Index (NDVI) - http://earthobservatory.nasa.gov/Features/MeasuringVegetation/measuring_vegetation_2.php.
- [8] Ortin, D. and Remondino, F., 2005. OCCLUSION-FREE IMAGE GENERATION FOR REALISTIC TEXTURE MAPPING, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Mestre-Venice, Italy.
- [9] SEII, E., 2002. Theory perceive the color, introduction to graphic of computer - <http://semmix.pl/color/default.htm>.
- [10] ZHANG, Y., Heipke, C., Butenuth, M. and HU, X., 2006. Automatic Extraction of Wind Erosion Obstacles by Integration of GIS Data, DSM and Stereo Images. *International Journal of Remote Sensing*, 27(8): 1677-1690.

Revised July 2011