Indirect building localization based on a prominent solid landmark from a forward-looking infrared imagery

Xiaoping Wang (汪小平)*, Tianxu Zhang (张天序), and Xiaoyu Yang (杨效余)

Institute for Pattern Recognition and Artificial Intelligence,

Huazhong University of Science and Technology, Wuhan 430074, China

**Corresponding author: wxphero2008@yahoo.cn*

Received November 11, 2010; accepted December 6, 2010; posted online March 28, 2011

A novel indirect building localization technique based on a prominent solid landmark from a forward-looking infrared imagery is proposed to localize low, deeply buried, or carefully camouflaged buildings in dense urban areas. First, the widely used effective methods are applied to detect and localize the solid landmark. The building target is then precisely indirectly localized by perspective transformation according to the imaging parameters and the space constraint relations between the building target and the solid landmark. Experimental results demonstrate this technique can indirectly localize buildings in dense urban areas effectively.

OCIS codes: 100.3008, 100.2000, 100.2960, 100.0100.

doi: 10.3788/COL201109.041003.

Building objects are often used as guides for aircraft navigation. However, the three-dimensional (3D) building localization is a difficult problem mainly because of the complexity of the scenes. Urban environments are very dense and composed of many types of buildings, making their analysis difficult. Thus, the reliable and efficient recognition of buildings is crucial for enabling such functionality. Moreover, building recognition from forward-looking infrared (FLIR) image sequences with a cluttered background is a challenging task. Traditionally, there are two modes to detect small targets in infrared (IR) images. One is the direct mode, which detects building targets directly. The other is the indirect mode, which first selects the prominent building objects as the solid landmark, and then detects the landmark, and finally localizes the targets. Generally, in the direct mode, most of the recent work in building detection has focused on the digital elevation models (DEMs), multiple aerial images, and generic models. The methods[1,2] based on DEMs are motivated by the fact that they already provide a geometric description of a scene from an aerial imagery or airborne scanner data. Multiple view aerial images are also the most common input data for building extraction[3−6] because such data are rich in terms of 3D information and allow the extraction of 3D primitives. Generic model-based building reconstruction techniques that combine several kinds of primitives are proposed by Baillard *et al.*[7] and Taillandier *et al.*[8], whereas the geometric model-based building recognition method is proposed by Yang *et al.*[9,10] for recognizing buildings from FLIR.

However, the DEMs, multiple aerial images, or generic model-based building extraction methods are always employed to extract visible or prominent buildings in downward-looking or forward-looking scenes. If the building is low, deeply buried, or carefully camouflaged in dense urban areas, the direct detection methods mentioned above are inefficient. Moreover, some building target enhancement techniques can be adopted to improve imaging quality and target detection probability in urban areas[11,12].

In this letter, which aims to overcome the disadvantages or enhance the performance of the direct mode approaches, an indirect mode approach based on a solid landmark is proposed to localize buildings in dense urban scenes. Experimental results show that the proposed method can largely improve the direct building detection performance of the existing widely used methods.

In dense urban scenes, if the targets are camouflaged or occluded by their surrounding objects, the direct navigation methods will not work. Therefore, indirect navigation techniques are adopted; that is, solid landmarks are selected to navigate the aircraft, as the landmark selection criterion is of utmost importance in aircraft navigation. According to the extensive experiments and detailed analysis, the selected solid landmarks should be prominent and practical to the ensure high localization probability of the aircraft. The criteria are as follows:

1) The solid landmark is not occluded or camouflaged by its peripheral solid object in the flying course of the aircraft.

2) The local contrast of the solid landmark is higher than its peripheral 3D solid object in the IR aerial image.

3) The real size of the solid landmark is larger or taller than its peripheral solid object in the flying course of the aircraft.

To localize the building target precisely in the IR image, the space constraint relations between the building target and the solid landmark in the actual scene, which

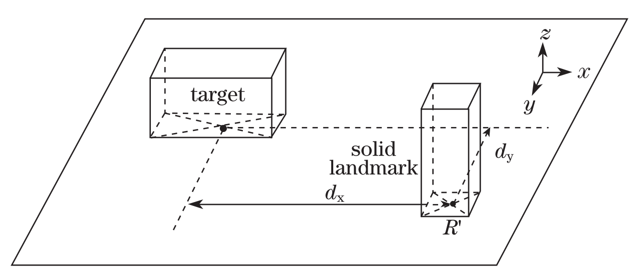

Fig. 1. Sketch map of the distance between the target and the solid landmark in DOM.

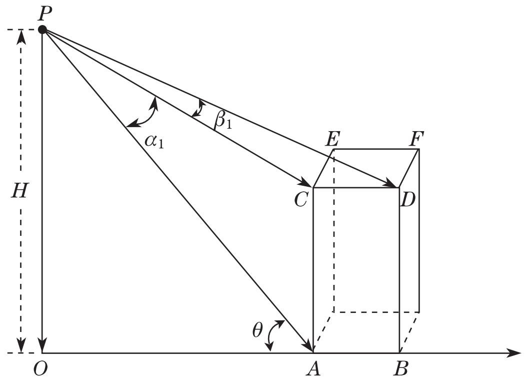

Fig. 2. Sketch map of the imaging in a section plane.

are denoted by d_x and d_y, respectively, should be obtained beforehand. Therefore, the space constraint relations, which are described in Fig. 1, can be expatiated in detail in the digital orthophoto map (DOM).

Let $R'(X'_r, Y'_r)$ and $T'(X'_t, Y'_t)$ denote the ground coordinates of the solid landmark and the building target in DOM. The space constraint relations $TP(d_x, d_y)$ between the solid landmark and the building target is calculated by $d_x = X'_t - X'_r$ and $d_y = Y'_t - Y'_r$.

According to the imaging parameters of the IR device and the actual size of the object, the imaging size of an object can be calculated by the following formulas. The sketch map of the imaging in a section plane is shown in Fig. 2.

Let H denote the imaging height. Let AC, CD, and DF denote the actual height, width, and length of building target, respectively. Let θ, φ, and ϕ denote the elevation, horizontal, and vertical angles of field of view, respectively. Let IMG_W and IMG_H denote the width and height of the IR image, respectively. Let OBJ_H, OBJ_W, and OBJ_L denote the imaging height, imaging width, and imaging length of target in IR image, respectively. We get

$$PA = H/\sin(\theta), \tag{1}$$

$$\angle PAC = \pi/2 - \theta, \tag{2}$$

$$PC^2 = PA^2 + AC^2 - 2 \times PA \times PC \times \cos(\angle PAC), \tag{3}$$

$$\angle\alpha_1 = \angle APC = \arccos\left(\frac{PA^2 + PC^2 - AC^2}{2 \times PA \times PC}\right), \tag{4}$$

$$\angle PCD = \pi + \angle\alpha_1 - \angle\theta, \tag{5}$$

$$PD^2 = PC^2 + CD^2 - 2 \times PC \times CD \times \cos(\angle PCD), \tag{6}$$

$$\angle\beta_1 = \angle CPD = \arccos\left(\frac{PC^2 + PD^2 - CD^2}{2 \times PC \times PD}\right), \tag{7}$$

$$\text{OBJ\_H} = \frac{\angle\alpha_1}{\angle\phi} \times \text{IMG\_H}, \tag{8}$$

$$\text{OBJ\_W}' = \frac{\angle\beta_1}{\angle\varphi} \times \text{IMG\_W}, \tag{9}$$

$$\text{OBJ\_W}'' = \frac{CD}{AC} \times \text{OBJ\_H}, \tag{10}$$

$$\text{OBJ\_W} = \min(\text{OBJ\_W}', \text{OBJ\_W}''), \tag{11}$$

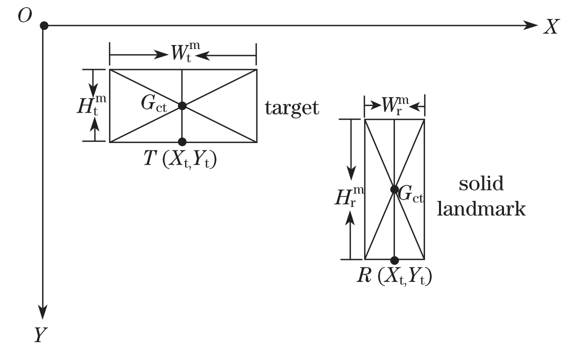

Fig. 3. Sketch map of the space constraint relations between the target and the solid landmark in image.

$$\text{OBJ\_L}' = \frac{DF}{AC} \times \text{OBJ\_H}, \tag{12}$$

$$\text{OBJ\_L}'' = \frac{DF}{CD} \times \text{OBJ\_W}, \tag{13}$$

$$\text{OBJ\_L} = \min(\text{OBJ\_L}', \text{OBJ\_L}''). \tag{14}$$

After the prominent solid landmark is detected successfully using the widely used methods, the building indirect localization stage begins. Let $G_{cr}(X_{cr}, Y_{cr})$ denote the centroid of the solid landmark in IR image. The imaging parameters of the IR device, the real size, and the imaging size of the building target and the landmark, and the constraint relations between the building target and the solid landmark in DOM are adopted to localize the target in the IR image.

Let $R(X_r, Y_r)$, $T(X_t, Y_t)$, and $O'(X_0, Y_0)$ denote the ground coordinates of the solid landmark, building target, and optical axis in the IR image, which are described in Fig. 3, where $G_{ct}(X_{ct}, Y_{ct})$, $G_{cr}(X_{cr}, Y_{cr})$, W_t^m, H_t^m, W_r^m, and H_r^m denote the centroid, imaging width, and imaging height of the target and the solid landmark, respectively. As described in Fig. 3, we know that $X_r = X_{cr}$, $Y_r = Y_{cr} + H_r^m/2$, $X_t = X_{ct}$, $Y_t = Y_{ct} + H_t^m/2$. The imaging size of the target and the solid landmark can be calculated by the formulas given above.

Considering that the optical axis aims at the center of the field of view in an actual imaging scene, then X_0 =IMG_W/2, Y_0=IMG_H/2. Let α denote the azimuth angle of the field of view. Let $\Delta\phi$ and $\Delta\varphi$ denote the vertical delta angle and horizontal delta angle between the landmark and optical axis, respectively. Let ΔX_{r2a} and ΔY_{r2a} denote the horizontal and vertical distances between the landmark and optical axis, respectively. Let ΔX_{t2a} and ΔY_{t2a} denote the horizontal and vertical distance between the target and optical axis, respectively. The sketch map of the perspective transformation is shown in Fig. 4. The related formulas are given as follows:

$$\begin{aligned} \Delta\phi &= \phi \times (Y_r - Y_0)/\text{IMG\_H}, \\ \Delta\varphi &= \varphi \times (X_r - X_0)/\text{IMG\_W}, \end{aligned} \tag{15}$$

$$\begin{aligned} x_1 &= \sqrt{y_1^2 + H^2} \times \tan(\Delta\varphi), \\ y_1 &= H/\tan(\theta + \Delta\phi), \end{aligned} \tag{16}$$

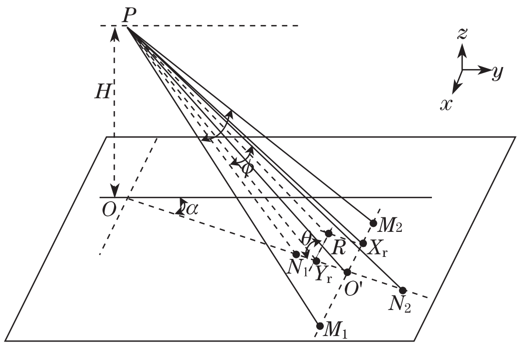

Fig. 4. Sketch map of the perspective transformation.

$$\Delta X_{\rm r2a} = x_1 \times \cos\alpha - y_1 \times \sin\alpha,$$
$$\Delta Y_{\rm r2a} = y_1 \times \cos\alpha + x_1 \times \sin\alpha, \quad (17)$$

$$\Delta X_{\rm t2a} = \Delta X_{\rm r2a} + d_{\rm x},$$
$$\Delta Y_{\rm t2a} = \Delta Y_{\rm r2a} + d_{\rm y}, \quad (18)$$

$$x_2 = \Delta X_{\rm t2a} \times \cos\alpha + \Delta Y_{\rm t2a} \times \sin\alpha,$$
$$y_2 = \Delta Y_{\rm t2a} \times \cos\alpha - \Delta X_{\rm t2a} \times \sin\alpha, \quad (19)$$

$$\theta_2 = \arctan(H/y_2),$$
$$\varphi_2 = \arctan(x_2/\sqrt{y^2 + H^2}), \quad (20)$$

$$X_t = X_0 + \varphi_2 \times \text{IMG\_W}/\varphi,$$
$$Y_t = Y_0 + (\theta_2 - \theta) \times \text{IMG\_H}/\phi. \quad (21)$$

Through calculation, the centroid of the buildings target, which is denoted by $G_{\rm ct}(X_{\rm ct}, Y_{\rm ct})$, can be localized precisely in the IR image, where $X_{\rm ct} = X_{\rm t}$, $Y_{\rm ct} = Y_{\rm t} - H_{\rm t}^{\rm m}/2$.

In this letter, the proposed method is used to detect the occluded 3D object in the IR image sequences. The IR image sequences were taken from a dense urban area in a high oblique view and gathered by a mid-wave IR sensor mounted to an aircraft under a variable depression angle and imaging distance.

Under the imaging parameters of the image sequence and the direction of the navigation as shown in Fig. 5, the building target is occluded by its surrounding three tall buildings. The combination of these three tall buildings, which are very prominent and are not occluded by its surrounding buildings in the imaging scene, is selected as the solid landmark. The DOM of the image size 800×800 (pixels) in the vicinity of the building target is presented in Fig. 5; thus, $X_{\rm t}' = 401$, $Y_{\rm t}' = 396$, $X_{\rm r}' = 522$, $Y_{\rm t}' = 285$, $d_{\rm x} = X_{\rm t}' - X_{\rm r}' = -121$, $d_{\rm y} = Y_{\rm t}' - Y_{\rm r}' = 111$. The real size of the three tall buildings is the same: 100 m in height and 40 m in width. The real size of the target is 100 m in height and 82 m in width. Thus, the imaging size of the solid landmark and building target can be calculated by the Eqs. (1)–(14).

The signal frame selected from the image sequences with a size of 320×256 (pixels) is presented in Fig. 6(a).

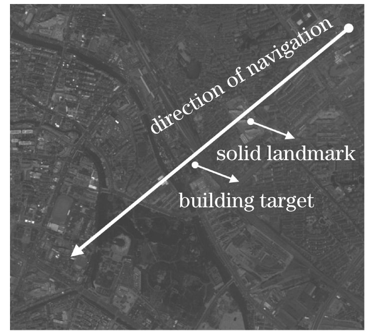

Fig. 5. Image of the DOM.

Based on the imaging conditions and the real size of the building target and the solid landmark, the imaging height is almost 20 pixels in Fig. 6(a), i.e., $H_{\rm t}^{\rm m} = 20$ and $H_{\rm r}^{\rm m} = 20$. Using the geometric model-based building detection method proposed by Yang *et al.*[9,10], the solid landmark can be successfully detected in an IR image. The solid landmark is localized in Fig. 6(b) with a white crossing, whose coordination is $X_{\rm cr} = 165$, $Y_{\rm cr} = 144$. Thus, $X_{\rm r} = X_{\rm cr} = 165$, $Y_{\rm r} = Y_{\rm cr} + H_{\rm r}^{\rm m}/2 = 164$. Based on the imaging parameters and the space constraint relations $TR(d_{\rm x}, d_{\rm y})$ between the solid landmark and the building target, the ground coordination of the building target $T(X_{\rm t}, Y_{\rm t})$ can be calculated by Eqs. (15)–(21). After calculation, $X_{\rm t} = 140$, $Y_{\rm t} = 138$. Finally, the building target, which is denoted by a white rectangle, is localized in Fig. 6(c). Its centroid is $X_{\rm ct} = X_{\rm t} = 140$, $Y_{\rm ct} = Y_{\rm t} - H_{\rm t}^{\rm m}/2 = 138 - 20 = 118$.

In conclusion, the proposed method has been applied to large-scale IR image sequences captured by a mid-wave IR sensor mounted to an aircraft under such variable imaging conditions as flying height, elevation, azimuth, and roll. Based on the detailed analysis of the experimental results, two suggestions are made to utilize the proposed method to improve the localization precision of a building target.

(1) The small errors of roll have several influences on the localization precision of the target localization procedure.

(2) If the errors caused by imaging conditions are not omitted, the flying height of the aircraft may be higher than 1 km or the elevation of the aircraft may be smaller than 5°.

The localization precision of IR image sequences, which are obtained under the guidance of the two suggestions, has an error of almost 5 pixels. Furthermore, the building detection performance has been improved largely compared with the traditional method. The detection probability has been maintained at higher than 90%.

(a) (b) (c)

Fig. 6. Mid-result images of the proposed method.

This work was supported by the National Natural Science Foundation of China (No. 60736010) and the Arm Pre-Research Key Foundation of China (No. 9140A01040309JW0505).

References

1. F. Lafarge, X. Descombes, J. Zerubia, and M. Pierrot-Deseilligny, ISPRS J. Photogram Rem. Sens. **63,** 365 (2008).
2. M. Ortner, X. Descombes, and J. Zerubia, Int. J. Comput. Vis. **72,** 107 (2007).
3. Z. Kim and R. Nevatia, Comput. Vis. Image Understand. **96,** 60 (2004).
4. P. Saeedi and H. Zwick, in *Proceedings of 10th Intl. Conf. of Control, Automation, Robotics and Vision* 623 (2008).
5. M. Fradkin, H. Maître, and M. Roux, Comput. Vis. Image Understand. **82,** 181 (2001).
6. S. Noronha and R. Nevatia, IEEE Trans. Pattern. Anal. Mach. Zntell. **23,** 501 (2001).
7. C. Baillard and A. Zisserman, in *Proceedings of the 19th ISPRS Congress* 33 (2000).
8. F. Taillandier and R. Deriche, in *Proceedings of ISPRS Congress* 35 (2004).
9. X. Yang, T. Zhang, and Y. Lu, J. Infrared Milli. Terah. Waves **30,** 468 (2009).
10. X. Yang, T. Zhang, and L. Yan, Proc. SPIE **7495,** 749531 (2009).
11. T. Yu, Q. Li, and J. Dai, Chin. Opt. Lett. **7,** 206 (2009)
12. S. Gao, C. Li, and D. Bi, Chin. Opt. Lett. **8,** 474 (2010)