Unsupervised Perception Model for UAVs Landing Target Detection and Recognition

Eric Bazan, Petr Dokládal, Eva Dokladalova

Unsupervised Perception Model for UAVs Landing Target Detection and Recognition

Eric Bazán[1], Petr Dokládal[1], and Eva Dokládalová[2]

[1] PSL Research University - MINES ParisTech, CMM - Center for Mathematical Morphology, Mathematics and Systems, 35, rue St. Honoré, 77305, Fontainebleau Cedex, France
{eric.bazan, petr.dokladal}@mines-paristech.fr
[2] IGM, Unité mixte de recherche CNRS-UMLV-ESIEE, UMR 8049, Université Paris-Est, Cité Descartes B.P.99, 93162, Noisy le Grand Cedex, France
eva.dokladalova@esiee.fr

Abstract. Today, unmanned aerial vehicles (UAV) play an interesting role in the so-called Industry 4.0. One of many problems studied by companies and research groups are the sensing of the environment intelligently. In this context, we tackle the problem of autonomous landing, and more precisely, the robust detection and recognition of a unique landing target in an outdoor environment. The challenge is how to deal with images under non-controlled light conditions impacted by shadows, change of scale, perspective, vibrations, noise, blur, among others. In this paper, we introduce a robust unsupervised model allowing to detect and recognize a target, in a perceptual-inspired manner, using the Gestalt principles of non-accidentalness and grouping. Our model extracts the landing target contours as outliers using the RX anomaly detector and computing proximity and a similarity measure. Finally, we show the use of error correction Hamming code to reduce the recognition errors.

Keywords: UAV, landing target, perception model, object detection, precision landing

1 Introduction

In this paper, we present a novel method for the detection of landing targets for the UAV vision aided landing. We propose to model the landing target by taking into account the principles of the human perception. The methodology presented works in an unsupervised mode, i.e., no need to adjust parameters.

In outdoor environments, many variables affect the vision-based landing target detection. The main problems to face are: the non-controlled light changes that generate shadowing, reflectance and saturation on the surfaces; the perspective and distance of the camera that deforms the objects; the motions and vibrations that blur the images and; the noise generation by a low-quality sensor.

The detection of the landing target can be viewed as an image segmentation problem, where there is a wide range of developed methods. The variational

framework [13], offers an optimal general method for image segmentation; however, its mathematical complexity and the constant selection of fidelity and a regularization parameters makes its use complex. Also, the number of iterations needed to find the optimal solution avoid having results in real-time. Conversely, thresholding methods have been used for the detection of landing targets [8][9] for its ease of use. However, for a good detection, its use is limited to indoor spaces, where the light conditions are controlled [1].

Recently, convolutional neural networks (CNN) techniques offer the possibility of detecting an object from a large set of classes with a high-reliability [3]. Nevertheless, these methods must have been trained with a database containing the object classes in a wide range of situations and, in case of changes in the object or the scene, the database must be rebuilt [19][6]. Besides, in some cases, the computation is carried out off-board the drone, which implies the need for network infrastructure and limitation of autonomy [10].

Humans can carry out the process of perception in a natural way [14]. We identify meaningful features and exciting events in a scene (such as points, lines, edges, textures, colors, movement) and with the help of our memory and the learning capacity we can recognize and classify objects. The primitives identification is a consequence of their non-accidental apparition, i.e., they are not generated randomly [2]. The Gestalt theory [17] states that we can build a whole (gestalt) through the grouping of non-accidental detected primitives. In this work, we explore the above ideas and propose a novel approach to detect a landing target in the same way as humans do, imitating the human perception process.

The work is organized as follows. In section 2 we develop the perception model. Specifically, subsection 2.2 describes how to retrieve image contours as meaningful primitives and subsection 2.3 describes how to group the contours to detect a landing target perceptually. Later, in section 3 we present the landing target design and the technique used for the methodology implementation; the results obtained are discussed in subsection 3.2. Finally, we present some conclusion and perspectives in section 4.

2 The Unsupervised Perception Model

2.1 Outset Presentation

The algorithm uses the contours as image primitives to obtain information about the scene. For the detection and recognition of landing targets, the algorithm is divided into three major stages. The first, localize all the image contours and extract meaningful contours using the non-accidentalness principle. The second stage computes some feature contours and, based on them, group the meaningful contours through the Gestalt laws. The last stage performs the target recognition using a decoding technique described in section 3. Figure 1 shows the three major stages and its subtasks.

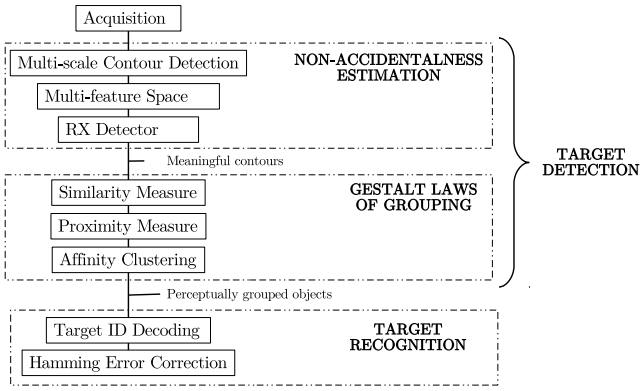


Fig. 1: Diagram of the phases for the landing target detection and recognition

2.2 Non-accidentalness Estimation

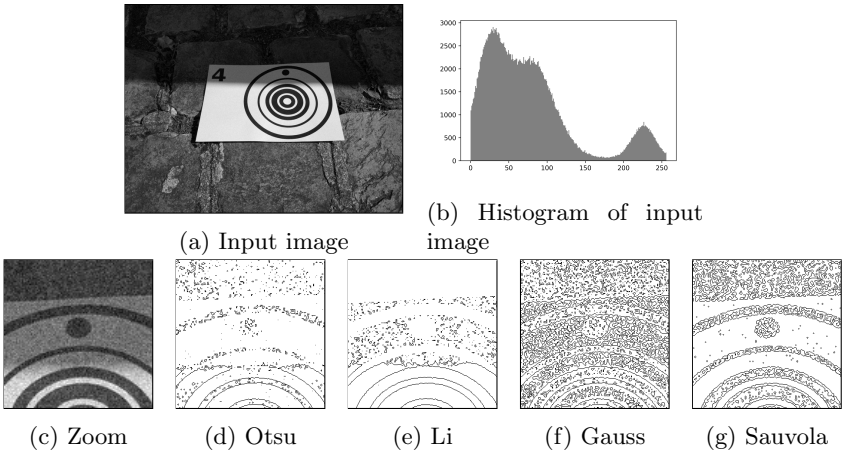
We aim to detect object contours in natural images where none, one or more landing targets can be present. Due to its real-time capacity, it is tempting to use a thresholding method to detect the contours of a binary image. We implemented several thresholding methods analyzed in [16], however, given the conditions where a landing target can be found, no method was found robust enough to variations in non-controlled outdoor environments. Figure 2a shows a landing target in an outdoor environment; we also show his histogram to highlight the levels of saturation in the scene. As a comparison, we take one representative method of each class of the taxonomy proposed in [16] to extract the contours of the image; clustering-based (fig. 2d), entropy-based (fig. 2e), spacial (fig. 2f) and local (fig. 2g) thresholding methods. Namely, there is no guarantee that the contours found by thresholding are present and continuous alongside the object borders.

Contour Detection Instead of using a thresholding method, we obtain the image without fixing any parameter. The use of the Marr-Hildreth [12] operator guarantees to find continuous and closed contours eliminating the possible noise in the image, while the contours of objects remain unchanged in the presence of shadows. This technique convolves the intensity image f with the 2-D Laplacian of Gaussian operator $\nabla^2 G(x, y, \sigma)$ and generates an image,

$$l_\sigma = \nabla^2 G(\sigma) * f \tag{1}$$

in which we localize the zero-crossings.

The parameter σ in eq. (1) permits to control the amount of image smoothing, but also acts as scale parameter, that when varies, it generates different scale-space images. Since no single filter can be optimal simultaneously at all scales [12], we use a multi-scale analysis [18] to detect the zero-crossings in l_σ at different scale-spaces to minimize the risk that some contour of interest is not detected.



(a) Input image (b) Histogram of input image

(c) Zoom (d) Otsu (e) Li (f) Gauss (g) Sauvola

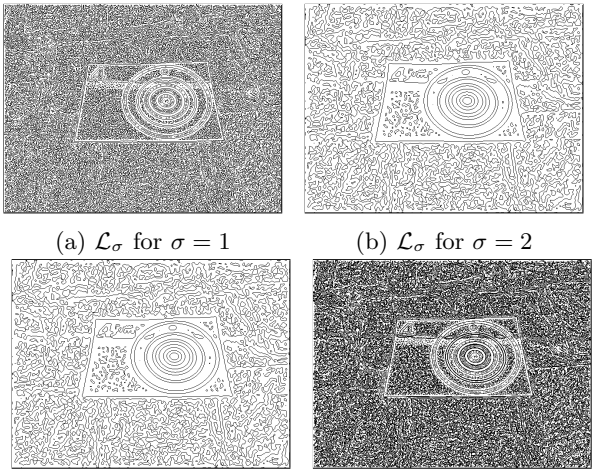
Fig. 2: Landing target under non-controlled illumination conditions and the controus obtained with some thresholding methods

The image l_σ from eq. (1) contains a set of contours $\mathcal{L}_\sigma = \{L_i^\sigma,\ \ i = 0, 1, \ldots, N\}$ for a given scale σ. Then,

$$\mathcal{L} = \bigcup_{\sigma} L_\sigma \tag{2}$$

represents all the contours of an image obtained at different scale-spaces. Figure 3d shows the set of contours $\mathcal{L}$ found for $\sigma = [1, 2, 3]$. Besides, it is also appreciated that at a fine scale (Fig. 3a) we can see more characteristics of the objects, i.e., there are more contours. Conversely, in coarse scales (Fig. 3c), due to the smoothing, there is a spatial distortion, and fewer contours appear. However, those contours that had already appeared at a coarse scale, will not disappear. Then, exist the probability that those contours that spatially coincide on two or more scales belong to a change of intensity generated by the border of an object.

Multi-feature Space The Helmholtz principle states that meaningful characteristics appear as large deviations from randomness and that is how the human perception automatically works to identify an object [2]. The a contrario model proposed in [4], formulates this principle statistically by setting the number of false alarms (NFA) below some acceptable level; however, this method cannot be easily extended to more complex shapes. Instead of setting the NFA, we use the RX detector [15] to detect outliers. Initially called the constant false alarms rate detection algorithm (CFAR) it can detect the presence of a know signal pattern in several signal-plus-noise channels. For that, it uses a $N \times Q$ multi-variable space $Z = [Z_1, \ldots, Z_Q]$ with Q observation vectors of dimension N. In our approach, the primitive is a closed contour. We build the multi-variable space with



(a) $\mathcal{L}_\sigma$ for $\sigma = 1$ (b) $\mathcal{L}_\sigma$ for $\sigma = 2$

(c) $\mathcal{L}_\sigma$ for $\sigma = 3$ (d) Set $\mathcal{L}$ for $\sigma = [1, 2, 3]$

Fig. 3: The image contours found at three different scales joined in the set $\mathcal{L}$

observations based on internal (geometrical features, e.g., circularity, roundness, area, perimeter) and external (e.g., mean gradient intensity, intensity inner area) properties of the contours.

Let $L_i \in \mathcal{L}$ be a contour, A_i its area and P_i its perimeter; we compute the circularity eq.(3) and the mean gradient intensity eq. (4) to build the multi-variable space $Z = [Z_1, Z_2]$.

$$Z_1 = \left[\frac{4\pi A_i}{P_i^2}, \;\; i = 0, \ldots, N \right]^T, \;\; N = card(\mathcal{L}) \tag{3}$$

$$Z_2 = \left[\frac{1}{P_i} \sum_{x \in L_i} \mid \nabla f(x) \mid, \;\; L_i \in \mathcal{L} \right]^T \tag{4}$$

RX Detector The RX anomaly detector [15] is commonly used to detect outliers on such data. The space Z models the set of contours $\mathcal{L}$ with $Q = 2$ feature vectors describing the circularity eq. (3) and the mean gradient intensity eq.(4). The RX detector gives an anomaly score to each contour taking into account the mean of the distribution and covariance between the Q-features through the Mahalanobis distance,

$$y_i = (z_i - \mu_Z)^T \Sigma_Z^{-1} (z_i - \mu_Z) \tag{5}$$

where $\mu_Z = [\mathrm{E}[z_1], \ldots, \mathrm{E}[z_N]]^T$ is the observations mean vector and Σ_Z^{-1} the $N \times Q$ covariance matrix of the data. If the data have normal random distribution,

then the score vector $Y = [y_i, \ldots, y_N]$ follows a chi-square distribution $\chi^2_Q(\varphi)$ with Q degrees of freedom, where φ is a confidence level [11]. The value of $\chi^2_Q(\varphi)$ with a confidence value $\varphi = 99.9\%$ operates as a threshold to identify all contours that behave as outliers in the multi-variable distribution. In our case, the contours belonging to a landing target appear as outliers in the vast majority of random contours belonging to the background.

With the previous strategy we preserve the anomalous contours having a value of mean gradient and circularity deviating from the principal mode of the distribution in the set $\widetilde{\mathcal{L}} = \{L_i \mid y_i > \chi^2_Q(\varphi)\}$. $\chi^2_Q(\varphi)$ is the value of the cumulative distribution at the confidence level φ and $\widetilde{\mathcal{L}} \subset \mathcal{L}$. It is essential to mention the importance of multi-scale contour detection of section 2.2; because it increases the number of samples in Z, allowing to build a richer multi-variable space.

In the set $\widetilde{\mathcal{L}}$ some contours make not part of a landing target. For example, in the figure 4, we can see that the paper sheet contours remain because they have a high value of circularity. The same occurs with the contours of those objects with an important value of mean gradient, as the number 4 at the top-left of the sheet or the rock textures of the background.

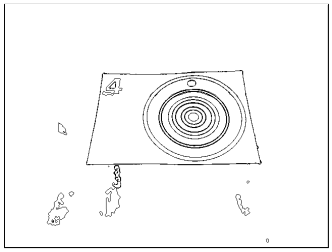
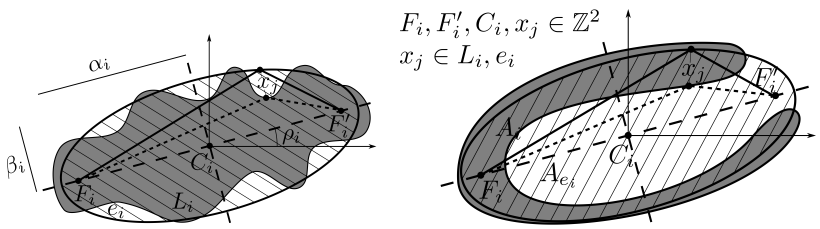


Fig. 4: The contours from Fig. 3d that behave as outliers in the multi-feature space Z with a confidence value of $\varphi = 99.9\%$

2.3 Gestalt Laws of Grouping

We use the Gestalt theory [17] to group the meaningful contours $L_i \in \widetilde{\mathcal{L}}$ and detect landing targets.

Goodness of Shape Since the landing targets have only circular contours, we evaluate the resemblance with an ellipse (to deal with the perspective deformation) of all contours. Considering an ellipse e_i that fits one gray contour L_i in Fig. 5a, we recover the centroid C_i, the rotational angle ρ, the semi-major axis α_i, the semi-minor axis β_i and the coordinates F_i and F'_i of the ellipse's foci. Then, the sum of the distances from any point of ellipse $x_j \in e_i$ to the foci is $\overline{x_j F_i} + \overline{x_j F'_i} = 2\alpha_i$. If the contour L_i is an ellipse, the value $d_i = \left|(\overline{x_j F_i} + \overline{x_j F'_i}) - 2\alpha_i\right|$ must be zero or negligible $\forall x_j \in L_i$.



(a) Affinity of a fit ω_i (b) Difference of area Δ_{A_i}

Fig. 5: Visual description of affinity of ellipse and difference of area

Based on the form of the landing target we estimate the the similarity using two measures,

$$\omega_i = \exp^{-\frac{d_i^2}{2\sigma^2}} \text{ the affinity of the fit and,} \tag{6}$$

$$\Delta_{A_i} = 1 - \frac{|A_{e_i} - A_i|}{\max(A_{e_i}, A_i)} \text{ the difference of area.} \tag{7}$$

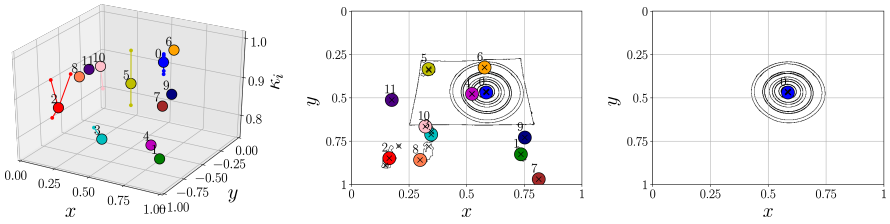
The affinity $\omega_i \to 1$ for contours closed to an ellipsoidal shape. However, if the contour L_i is a croissant shape (as in fig. 5b) then, the eq. (6) also has a high value (near to 1) but the contour is from being an ellipse. The variable in eq. (7) complements the affinity ω_i taking into account the area of the ellipse A_{e_i} and the area of the contour A_i. To calculate the similarity to an ellipse, we use the harmonic mean of both.

$$\kappa_i = \mathcal{H}(\omega_i, \Delta_{A_i}), \;\; \kappa_i \in (0, 1) \tag{8}$$

where $\kappa_i \to 1$ for contours ressembling to an ellipse and $\kappa_i \to 0$ otherwise. $\mathcal{H}$ denotes the harmonic mean $\mathcal{H} = N \left(\sum_{i=1}^{N} \xi_i^{-1} \right)^{-1}$.

Proximity Measure The Gestalt law of proximity states that we group those meaningful elements if they are spatially close to each other. In the case of contours, we take the coordinates of their centers C_i to measure their spatial proximity.

Affinity Clustering The normalized coordinates of the centroid C_i and the ellipse similarity κ_i map the contour $L_i \in \widetilde{\mathcal{L}}$ into the 3-D space $(0, 1) \in \mathbb{R}^3$. We use the affinity propagation clustering method [5] to group the contours using the matrix $X = [C_i, \kappa_i]$. This technique yields a set of clusters $\mathcal{C}_K \in \mathcal{C}(X)$. Because the landing target has ten different contours (see section 3), the clusters with $card(\mathcal{C}_K) \geq 10$ and an important similarity value $\mathcal{H}(\kappa_i) \geq 0.8$, represent the candidate contours of a landing target.



(a) Clusters obtained by affinity propagation

(b) Clusters projected on the image domain

(c) Target candidate cluster

Fig. 6: Clusters of contour from Fig. 4

The affinity propagation technique groups in $K = 12$ clusters the image contours from figure 4. In a 3D plot (fig. 6a), we see the influence of κ_i at clustering process. Projecting the clusters in a 2-D plane (fig. 6b), we notice that even if the contours are nearby, it can form a new cluster if there is a distant κ. A clear example is the clusters 0 and 4 (blue and purple, respectively) that correspond to the contour centers of the landing target and the center of the sheet of paper, they are close to each other but the similarity not. Applying the threshold values $card(\mathcal{C}_K) \geq 10$ and $\mathcal{H}(\kappa_i) \geq 0.8$ we obtain the candidate clusters to form a landing target (see fig. 6c).

Heretofore, we have built a model based on perceptual characteristics for the landing target detection. However, there could be false detections if there are circular objects with concentric borders in the image. We code an ID number in the target design to differentiate a landing target from an object with concentric circular edges. The coding of information allows discriminating between several landing targets and circular objects. The following section describes the landing target design as well as the coding and decoding technique.

3 Implementation

3.1 Landing target description

The landing target is formed by a set of black and white circles (see Fig. 7) that generate contours when stacked. Two of the circles ($\o_9$ and $\o_{10}$) have a constant diameter and form the ring that defines the target. The black circle ($\o_{11}$) is an orientation reference and has the same diameter as the smallest circle, $\o_{11} = \o_1$. The other circles $\o_1, \ldots, \o_8$ are coding circles.

Landing Target ID Encoding Let $\O = (\o_1, \o_2, \ldots, \o_n)$ denote the nominal diameters of the coding circles. We can set the nominal diameters, e.g., $\o_i = \frac{i}{n}\o_n$ for a target without the encoding capability. To encode a number in the target form, we modify the nominal diameters $\O$ to obtain $\O' = \left(\o'_1, \o'_2, \ldots, \o'_n\right)$ by

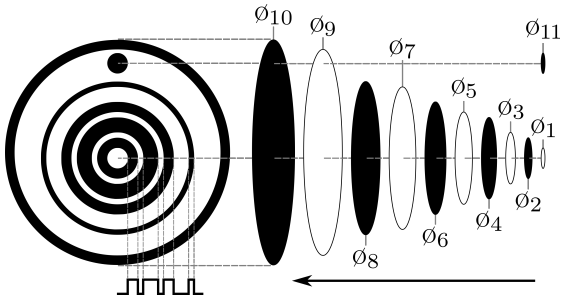


Fig. 7: Landing target design and description

adding/subtracting a positive constant Δh

$$\phi_i' = \begin{cases} \phi_i + \Delta h, & \text{if } w_i = 1 \\ \phi_i - \Delta h, & \text{otherwise} \end{cases} \tag{9}$$

and obtain a binary message $W = [w_1, \ldots, w_n]$. The message W is protected from errors by Hamming error-correction code [7]. It provides a set of different codewords $W = D \times M$ of size $n = k+m$, where D is useful data, $M = [I_k \mid 1-I_k]$ the generator matrix and I_k is the $k \times k$ the identity matrix. The data vector D comes from the decimal to binary conversion of the landing target ID* number. In our representation, we have experimented with $n = 8$ coding circles allowing to have four rings and $n = 8$ contours $\phi_1, \ldots, \phi_8$. This allows us to use the extended $[n, k]$ Hamming code with $k = 4$ data bits and $m = 4$ parity bits to generate $2^4 = 16$ landing targets.

Landing Target ID Decoding After the clustering stage of section 2.3, we rank by size the ellipses' major axes α_i by size and normalize them w.r.t. the largest value α_{10} to obtain $\boldsymbol{\alpha} = \frac{\phi_{10}}{\alpha_{10}}(\alpha_1, \ldots, \alpha_{10})$

We compare the received and normalized axis $\boldsymbol{\alpha}$ with the nominal diameters of the coding circles Ø and transform them into a binary vector $\widehat{W}$;

$$\widehat{W} = \begin{cases} 1, & \text{if } \alpha_i - \phi_i > 0 \\ 0, & \text{otherwise} \end{cases} \quad \forall i = 1, \ldots, n \tag{10}$$

The Hamming syndrome vector $S = \widehat{W} \times H^T$ (with $H = [1 - I_k \mid I_k]$ as the parity-check matrix) indicates whether an error has occurred. The syndrome is a null vector $S = 0$ when no error has occurred, otherwise, $S \neq 0$ and $\widehat{W} = W + E$. The element $e_i = 0$ of the error vector $E = H^T - S$ indicates an error at the position i. The $[8, 4]$ Hamming code can find up to two erroneous bits and correct one. Once the algorithm corrects the error (if there is), the vector $\widehat{W}$ is decoded by using the modulo 2 of the product $\widehat{D} = \widehat{W} \times M^T$.

*Identification number

3.2 Validation and tests

The presented strategy was validated on landing target images under simulated and real situations. We tested the algorithm in a synthetic image database which simulates four image degradations: noise, shadows, target deformation and change of size. For the real situations, we carried out several tests in indoor and outdoor scenarios. Figure 8 shows three interesting experiments and the output image of each stage of section 2.

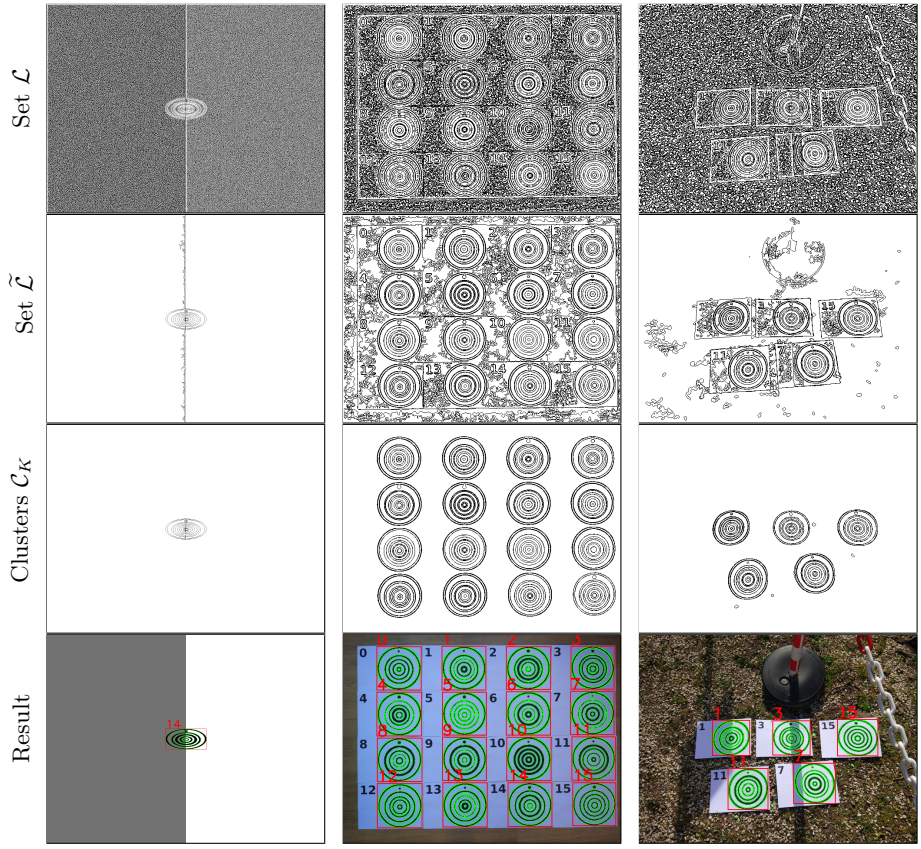


Fig. 8: Algorithm validation: (a) Target under simulated degradations, (b) The 16 targets in an indoor environment, (c) Five targets in an outdoor scenario under non-controlled image degradations

The first experiment (Fig. 8a) shows the four synthetic degradations together on landing target ID 14. In this context, the synthetic image represents the values of degradation maximum that the algorithm supports. Second experiment (Fig.

8b) was done in an indoor space to show the sixteen possible landing targets. In the scene there are no other objects. Finally, the last experiment (Fig. 8c) shows five landing targets in a more complex outdoor environment. Notice the presence of other objects, different background textures, irregular shadows and perspective deformation and change of scale of the landing targets.

In the three experiments, i) the non-accidentalness estimation stage eliminates the contours generated by noise with low circularity and mean gradient values; ii) the grouping stage filters random contours generated by intensity changes like shadows to keep contours with an important value of similarity and proximity. The compilation of the experiments carried out under real conditions can be seen in https://youtu.be/igsQc7VEF2c.

4 Conclusion and Future Extensions

We have described the procedure for the landing target detection and recognition based on a perception model. The algorithm is based on the Helmholtz non-accidentalness principle and the Gestalt theory. The non-accidentalness estimation is performed in a multi-feature object space built from the image contours at different scales. This approach allows us to obtain scene information avoiding the loss of information because of the objects' change of size or the presence of shadows and noise. We have used the similarity and proximity Gestalt laws to group the contours and build a perceptual object and the Hamming error codes to perform the landing target recognition. The experiments show that the proposed methodology for the detection of landing targets is robust to uncontrolled light conditions and other images degradations existing in complex environments.

For this particular work, the objective was to detect a landing target, a circular object (basic geometric shape) with concentric borders; however, the presented methodology has a wide extension capacity. At the moment we use only the image contours to measure the circularity and the mean intensity gradient. A future extension could include the use of other image primitives, such as points, regions, texture or color, and other object features. Similarly with the grouping laws, we explored only the proximity and similarity laws; however, it is possible to use other laws such as alignment, symmetry or continuity. These ideas will allow creating new descriptive feature spaces of scenes and objects in an image; with the possibility of make combinations between grouping laws, features and image primitives to detect some particular object.

Finally, our contribution proposes a perception model that utilizes the a contrario theory but avoids the mathematical complexity of setting the NFA to adjust threshold values or parameters in specific situations.

Acknowledgments

This research is partially supported by the Mexican National Council for Science and Technology (CONACYT).

References

1. Araar, O., Aouf, N., Vitanov, I.: Vision Based Autonomous Landing of Multirotor UAV on Moving Platform. Journal of Intelligent & Robotic Systems 85(2), 369–384 (Feb 2017)
2. Attneave, F.: Some informational aspects of visual perception. Psychological Review 61(3), 183–193 (1954)
3. Carrio, A., Sampedro, C., Rodriguez-Ramos, A., Cervera, P.C.: A Review of Deep Learning Methods and Applications for Unmanned Aerial Vehicles. J. Sensors 2017, 3296874:1–3296874:13 (2017)
4. Desolneux, A., Moisan, L., Morel, J.M.: From Gestalt Theory to Image Analysis: A Probabilistic Approach. Interdisciplinary Applied Mathematics, Springer-Verlag, New York (2008)
5. Frey, B.J., Dueck, D.: Clustering by passing messages between data points. Science (New York, N.Y.) 315(5814), 972–976 (Feb 2017)
6. Furukawa, H.: Deep Learning for End-to-End Automatic Target Recognition from Synthetic Aperture Radar Imagery. arXiv:1801.08558 [cs] (Jan 2018)
7. Hamming, R.W.: Error detecting and error correcting codes. The Bell System Technical Journal 29(2), 147–160 (Apr 1950)
8. Lacroix, S., Caballero, F.: Autonomous detection of safe landing areas for an UAV from monocular images. In: In IEEE/RSJ International Conference on Intelligent Robots and Systems (2006)
9. Lange, S., Sünderhauf, N., Protzel, P.: Autonomous Landing for a Multirotor UAV Using Vision. In: In SIMPAR 2008 Intl. Conf. on Simulation, Modeling and Programming for Autonomous Robots. pp. 482–491 (2008)
10. Lee, J., Wang, J., Crandall, D., Šabanović, S., Fox, G.: Real-Time, Cloud-Based Object Detection for Unmanned Aerial Vehicles. In: 2017 First IEEE International Conference on Robotic Computing (IRC). pp. 36–43 (Apr 2017)
11. Lu, C.T., Chen, D., Kou, Y.: Multivariate spatial outlier detection. International Journal on Artificial Intelligence Tools 13(04), 801–811 (Dec 2004)
12. Marr, D., Hildreth, E.: Theory of edge detection. Proc. R. Soc. Lond. B 207(1167), 187–217 (Feb 1980)
13. Mumford, D., Shah, J.: Optimal approximations by piecewise smooth functions and associated variational problems. Communications on Pure and Applied Mathematics 42(5), 577–685 (Jul 1989)
14. Petitot, J.: Neurogéométrie de la vision: modèles mathématiques et physiques des architectures fonctionnelles. Editions Ecole Polytechnique (2008)
15. Reed, I.S., Yu, X.: Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. IEEE Transactions on Acoustics, Speech, and Signal Processing 38(10), 1760–1770 (Oct 1990)
16. Sezgin, M., Sankur, B.: Survey over image thresholding techniques and quantitative performance evaluation. J. Electronic Imaging 13(1), 146–168 (Jul 2010)
17. Wertheimer, M.: FormsUntersuchungen zur Lehre von der Gestalt II. Psycologische Forschung 4, 301–350 (1923)
18. Witkin, A.: Scale-space filtering: A new approach to multi-scale description. In: ICASSP '84. IEEE International Conference on Acoustics, Speech, and Signal Processing. vol. 9, pp. 150–153 (Mar 1984)
19. Yao, H., Yu, Q., Xing, X., He, F., Ma, J.: Deep-learning-based moving target detection for unmanned air vehicles. In: 2017 36th Chinese Control Conference (CCC). pp. 11459–11463 (Jul 2017)