Fast Automatic Detection of Wildlife in Images from Trap Cameras*

Karina Figueroa[1], Antonio Camarena-Ibarrola[2],
Jonathan García[1], and Héctor Tejeda Villela[1]

[1] Facultad de Ciencias Físico-Matemáticas,Universidad Michoacana, México
[2] Facultad de Ingeniería Eléctrica, Universidad Michoacana, México
{karina,htejeda}@fismat.umich.mx, camarena@umich.mx

Abstract. Photo-identification of naturally marked animals is a non-intrusive technique for obtaining valuable information regarding population size and behavior in the wilderness of endangered species. In this paper we present a method for detecting/cutting wild felines in pictures taken with trap-cameras installed in the forest and triggered by infrared sensors. The detection of these felines serves the purpose of collecting information useful in studies about the population size or the migration phenomena. We propose computing the difference of images from the same trap-cameras within a short period of time. According to our experiments, our method is fast, reliable and robust, this method can be used for other species with different pelage patterns.

1 Introduction

Environmentalists are interested in helping species in danger to be extinct and they need a way to monitor the population of such species. An intrusive way to accomplish it consists of capturing each individual and mark it artificially. Applying artificial marks (i.e. Radio-Frequency Identificators)) to wild animals may alter their normal behavior and therefore reduce the effectiveness of the method [9], such solution requires a high amount of resources both human and material and so it is frequently unaffordable. The marking process itself may be disruptive due to the need of handling and restraining the wild animals. Natural markings of wild animals may be used instead [8,6].

A clever non intrusive solution consists of installing trap-cameras which are triggered by infrared sensors, whenever an animal crosses in front of the camera a picture is taken and stored for future processing. These pictures are processed by hand normally to identify individual members of endangered species using expert and skilled people for that purpose [11]. Using people for such task is both expensive, slow, and not very accurate due to human errors. We are interesting in a method that would allow to automatically identify wild felines for the purpose of taking a census.

* This work is partially funded by National Council of Science and Technology (CONACyT) of México and Universidad Michoacana de San Nicolás de Hidalgo, México.

E. Bayro-Corrochano and E. Hancock (Eds.): CIARP 2014, LNCS 8827, pp. 940–947, 2014.

Foster & Harmsen [5] performed a review of 47 published studies. They questioned the validity of the population estimates from camera-trap data since in these studies the members of the population were identified by phenotypic and environmental induced characteristics like scars, broken nails, dermal parasites, hair loss, etc., in general, features that change with time. Recognition of individuals using such characteristics produces misidentifications which in turn induce errors in population estimates. It makes sense to use image processing techniques to identify individuals from select species with unique coat patterns (e.g., spots or stripes) to perform census automatically.

Wildlife images have recently been used to implement photo-identification systems for naturally marked animals [8,6]. In [8] cheetahs, from the Serengeti National Park in Tanzania, were recognized semi-automatically extracting by hand the area between the shoulder blade, the hip joint, the belly line, and the backbone, afterwards a sample of the coat pattern is extracted for each animal and stored in the computer as a matrix of numbers, consisting of gray-scale intensities, they call these features, the Identifier Array (IA). The correlation coefficient between different IAs taken from different cheetahs was used as a similarity measure. This technique is not very efficient in terms of processing time and not fully automatic since it requires the work of an operator to process each image.

Bolger et al [3] worked in identification of wild animals based on coat patterns for subsequent photographic mark-recapture analysis, they used SIFT points for individual animal identification for strongly marked texture species. SIFT stands for Scale Invariant Feature Transform and it is an algorithm used in Computer Vision (CV) to detect local points of interest using a histogram of gradients of the image [7] .

Yu *et al* [12] combined SIFT and cLBP descriptors (cell-structured Local Binary Pattern) [13] aiming for robust characterization under both linear and nonlinear luminance variation. Yu *et al* claimed that SIFT and cLBP combined outperform SIFT alone in the task of animal identification using trap cameras.

Recently, a software system called IS3 was designed as an aid in the process of identifying some sort of individuals [1,10]. IS3 extracts all key points within a region of interest as indicated by the user. Each key point has a size and a location. The image should be taken ideally perpendicular to the line of sight and no more than 30 degrees.

We are interested in pattern recognition of Ocelots in Michoacán, and the first problem to solve here is the automatic segmentation of the Region of Interest (ROI) (i.e. The ocelot).

2 Data Used

Charre-Medel *et al* in [4] showed the presence of jaguars (Panthera onca) in the state of Michoacán, México back in 2010 and documented it with six photographs obtained with camera-traps and a skull collected in the field. These records belong to the tropical semi-deciduous forest in a transitional area between *Sierra*

Madre del Sur and the Pacific Coast. In particular, they also have about 300 records of Ocelots in which we are interested.

Our proposal is based on images captured in the open field using trap-cameras in places previously selected by experts in the field. Cameras are fixed (not moving), which is an important fact for us. Comparing two images captured consecutively by the same camera, the background may stay unchanged, while the wild animal will change due to its motion. It is important to mention that images can be captured during the whole day. At night, trap-cameras take photos using flash.

Figure 1 shows two consecutive photos with a difference of 8 seconds. Observe the variation in shadows and vegetation.

Fig. 1. Images from a trap-camera. Sample 1

3 Proposed Method

In order to segment the images we present two methods, an intuitive method (Diff of images) and the other makes use of filters.

3.1 Diff of Images

Our proposal is based in the following hypothesis, two images taken with the same camera within a very short period of time would differ only by the presence of wild animals, with this in mind, we compute the difference between the two images. There are two ways of doing this:

- Take the difference between the two images *pixel by pixel*
- Take the difference between the two images using a *squared window* of size m, see Algorithm 1

Pixel by Pixel's Method. For this method we compare pixel by pixel between two different images (f and t). That is,

$$g(x,y) = \begin{cases} 255 & \text{if } f(x,y) - t(x,y) \geq \alpha \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

We define threshold α (a free parameter) to control the difference between images, it might be the case that some moving body appears in one image (i.e. an animal).

Squared Window's Method. In this method we are using the value at a point in the context of a region. The hypothesis of this method is that in case of a change of shade or light on stage, they are not reflected only in a pixel but in a whole region, therefore, a region illuminated or shady would have the same value with respect to its center.

In Algorithm 1 we use β to decide which pair of images can be considered as the first one having the presence of the ocelot and the second one not (presence-absence) or if we should instead consider the first image as not having an ocelot and the second having it (absence-presence). In the same algorithm Distance-SubmatrixToCentre equals $\sum d(I(i+r, j+c), I(i,j))$ where r and c are the rows and columns respectively that cover a submatrix.

Algorithm 1. diff based in a submatrix(f,t) ; Let f and t be images in RGB

OUTPUT: 2 new images I_f, I_t with the diff.
Let α a threshold permitted, and $d_1^{'} \leftarrow 0$,
for $i \leftarrow 1$ to $n-m$ **do**
  for $j \leftarrow 1$ to $n-m$ **do**
    $d_1 = DistanceSubmatrixToCentre(f,i,j)$
    $d_2 = DistanceSubmatrixToCentre(t,i,j)$
    $P_1 \leftarrow$ WHITE, $P_2 \leftarrow$ WHITE
    if $|d_1 - d_2| > \alpha$ **then**
      if $|d_1^{'} - d_1| > \beta$ **then**
        $P_1 \leftarrow$ BLACK
      else
        $P_2 \leftarrow$ BLACK
      end if
      $I_f[i,j] \leftarrow P_1$ and $I_t[i,j] \leftarrow P_2$
      $d_1^{'} \leftarrow d_1$
    end if
  end for
end for

3.2 Classical Filters

We tried segmentation based on textures, with an entropy filter we can identify regions of different textures, Our hypothesis is that the texture of the image in the region where an ocelot is present differs from the texture of the rest of the picture formed by rocks, trees, etc. After applying the entropy filter we convert the resulting gray scale image to b/w using a threshold of 0.93.

4 Experiments and Results

We will discuss the results obtained after using the algorithms explained above.

4.1 Pixel by Pixel

As we expected we did not obtain good results by computing the difference pixel by pixel. We had many problems, first, we did not know whether the difference obtained was for the case of presence-absence or for the absence-presence case. Second, this strategy is too sensitive to changes in projected shadows, movement of leaves caused by the wind, and of course light changes.

4.2 Squared Window

Using the second strategy, we show how the performance of this technique is affected by parameters m and α.

4.3 Studying Parameters

m size. In Figure 2 we show how parameter m affects the segmentation of the ocelot. Notice that when $m = 3$ the result is very good, while for $m = 5$ or $m = 7$ the method is too sensitive to trivial differences between the images.

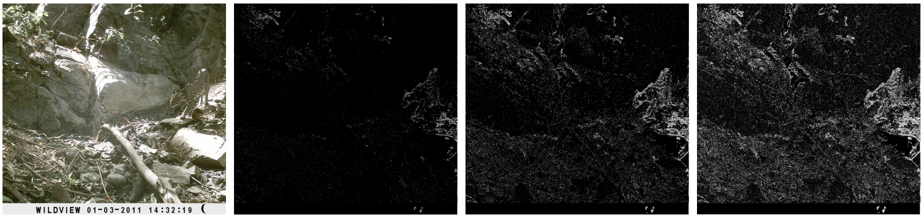

Fig. 2. Changing m size, from left to right original, m =3,5,7

α's Values. In Figure 3 we show how α affects the sensitive of our method as expected . Notice that if α is high it barely detects something in the images.

4.4 Presence-Absence

In figure 4 we show the hardest problem with this technique. When we detect a difference between both images we do not know if it is a case of presence-absence or absence-presence. We propose the use of a third parameter (β) to solve this problem. With $\beta = 3$ we get these images where we can see both animals in a single image.

4.5 Best Results

Figure 5 shows a very complicated scenario. The best results were obtained with our technique. Observe how we could easily segment the animal from these resulting images. In particular Figures 5 at the bottom we obtained excellent results when comparing with other techniques.

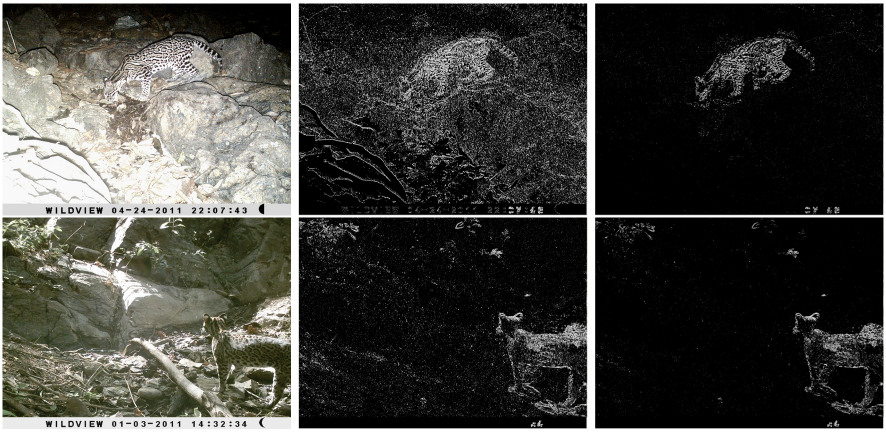

Fig. 3. Changing α. Parameters used from left to right, top: original, α =16, and 32. Bottom: original, α =8 and 16

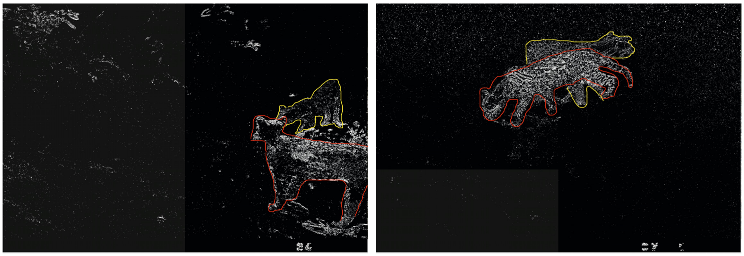

Fig. 4. Showing the problem of presence-absense in some images. Lines in yellow and red were marked manually

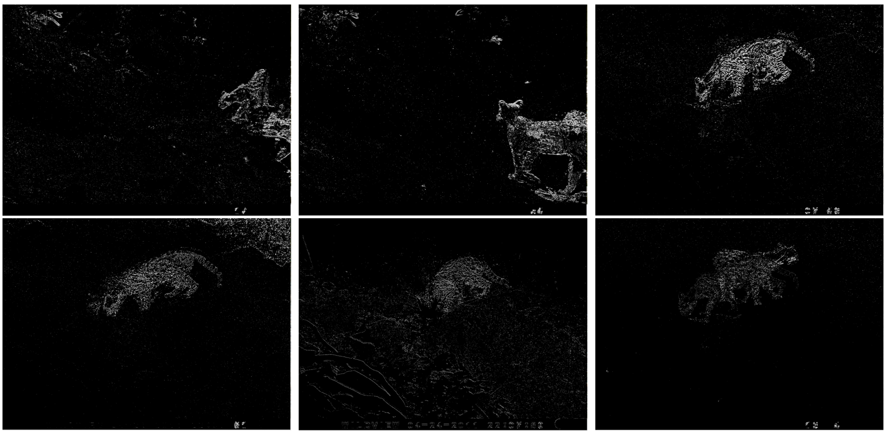

Fig. 5. Best results using the follow parameters, from left to right, top: $m = 5, \alpha = 8$, $m = 7, \alpha = 16$ and $m = 7, \alpha = 32$; bottom: $m = 7, \alpha = 32$, $m = 3, \alpha = 16$, and $m = 5, \alpha = 16$

4.6 An Alternative Solution

The Speeded Up Robust Features (SURF) algorithm [2] finds a great number of points in the region of the some images where ocelots are present, this fact could be used for segmenting the ocelot, see Figure 6. However, our method segments the region where an ocelot is present in much less time and it works for all the images in different scenarios.

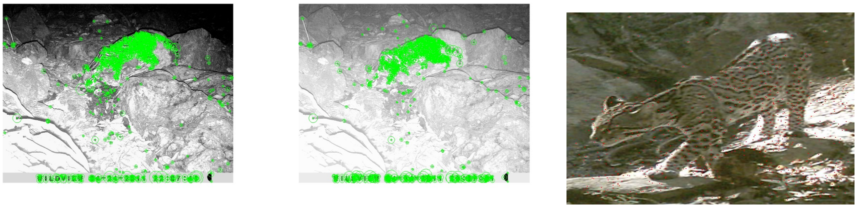

Fig. 6. Images computed with SURF. The third one is cutted but it shows how bad can works surf technique in some of our images.

5 Conclusions and Future Work

In this paper we proposed a method for detection of ocelots in pictures taken with trap-cameras installed at several locations in the forest. According to our experiments, our method is fast, reliable and robust, this method can be used for other species that also have complex and variable pelage patterns. The next stage of this research consists in building the recognizer to identify individual ocelots and not just the presence of one of them.

Acknowledgements. We are thankful with Dr. Tiberio Monterrubio for sharing his collection of photos with us.

References

1. Van Tienhoven, A.M., Den Hartog, J.E., Reijns, R., Peddemors, V.M.: A computer-aided program for pattern-matching of natural marks of the spotted raggedtooth shark carcharias taurus (rafinesque, 1810). Journal of Applied Ecology (2007) (in press)
2. Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: Surf: Speeded up robust features. Computer Vision and Image Understanding (CVIU) 110(3), 346–359 (2008)
3. Bolger, D.T., Morrison, T.A., Vance, B., Lee, D., Farid, H.: A computer-assisted system for photographic mark–recapture analysis. Methods in Ecology and Evolution 3(5), 813–822 (2012)
4. Charre-Medellín, J.F., Monterrubio-Rico, T.C., Botello, F.J., León-Paniagua, L., Nez, R.N.: First records of jaguar (panthera onca) from the state of michoacán, méxico. BioOne Researcch Evolved 58(2), 264–268 (2013)

5. Foster, R.J., Harmen, B.J.: A critique of density estimation from camera-trap data. The Journal of Wildlife Management (2011)
6. Whitehead, H., Christal, J., Tyack, P.L.: Studying cetacean social structure in space and time. In: Cetacean Societies: Field Studies of Dolphins and Whales (2000)
7. Ke, Y., Sukthankar, R.: Pca-sift: A more distinctive representation for local image descriptors. In: Computer Vision and Pattern Recognition, pp. 511–517 (2004)
8. Kelly, M.J.: Computer-aided photograph matching in studies using individual identification: An example from serengeti cheetahs. Journal of Mammalogy 82(2), 440–449 (2001), http://asmjournals.org/doi/abs/10.1644/1545-1542%282001%29082%3C0440%3ACAPMIS%3E2.0.CO%3B2
9. Gauthier-Clerc, M., Gendner, J.-P., Ribic, C.A., Fraser, W.R., Woehler, E.J., Descamps, S., Gilly, C., Le Bohec, C., Le Maho, Y.: Long-term effects of flipper bands on penguins. Proc. R Soc. Lond. Ser. B-Biol. Science 271, 423–426 (2004)
10. Team, I.: Interactive individual identification software. webpage (April 2014), http://www.reijns.com/i3s
11. Tempa, T., Norbu, N.: P., D., Nidup, T.: Results from A Camera Trapping Exercise for Estimating Tiger Population Size in the Lower Foothills of Royal Manas National Park. KUENSEL Corporation Ltd. (2011)
12. Yu, X., Wang, J., Kays, R., Jansen, P.A., Wang, T., Huang, T.S.: Automated identification of animal species in camera trap images. EURASIP Journal on Image and Video Processing (2013)
13. Zhang, J., Huang, K., Yu, Y., Tan, T.: Boosted local structured hog-lbp for object localization. In: Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2011, pp. 1393–1400. IEEE Computer Society, Washington, DC (2011), http://dx.doi.org/10.1109/CVPR.2011.5995678