JOURNAL OF CRITICAL REVIEWS

ISSN- 2394-5125 VOL 06, ISSUE 01, 2019

Detection algorithms that extract moving objects from scenes with cluttered backgrounds

Gosukonda Lalitha[1], Ronda Ravindra Reddy[2]

Assistant Professor[1,2]
Department of ECE
MREC

Abstract- The advantages of using color as feature to achieve object's similarity is analyzed and found that it is robust against the complex, deformed and changeable shape (i.e. different human profiles). In addition, it is also scale and rotation invariant, as well as faster in terms of processing time. Color information is extracted, stored and compared to find uniqueness of each object.
Thus far, we have mainly deal with hardware and software setting up. After creating working environment, we have utilized some detection algorithms that extract moving objects from scenes with cluttered backgrounds.
Index Terms- Extract, Detection, cluttered Background, Object

I. INTRODUCTION

An image is a two-dimensional picture, which has a similar appearance to some subject usually a physical object or a person. Image is a two-dimensional, such as a photograph, screen display, and as well as a three-dimensional, such as a statue. They may be captured by optical devices—such as cameras, mirrors, lenses, telescopes, microscopes, etc. and natural objects and phenomena, such as the human eye or water surfaces. The word image is also used in the broader sense of any two-dimensional figure such as a map, a graph, a pie chart, or an abstract painting. In this wider sense, images can also be rendered manually, such as by drawing, painting, carving, rendered automatically by printing or computer graphics technology, or developed by a combination of methods, especially in a pseudo-photograph.

Fig 1:Image

An image is a rectangular grid of pixels. It has a definite height and a definite width counted in pixels. Each pixel is square and has a fixed size on a given display. However different computer monitors may use different sized pixels. The pixels that constitute an image are ordered as a grid (columns and rows); each pixel consists of numbers representing magnitudes of brightness and color.

ISSN- 2394-5125 VOL 06, ISSUE 01, 2019

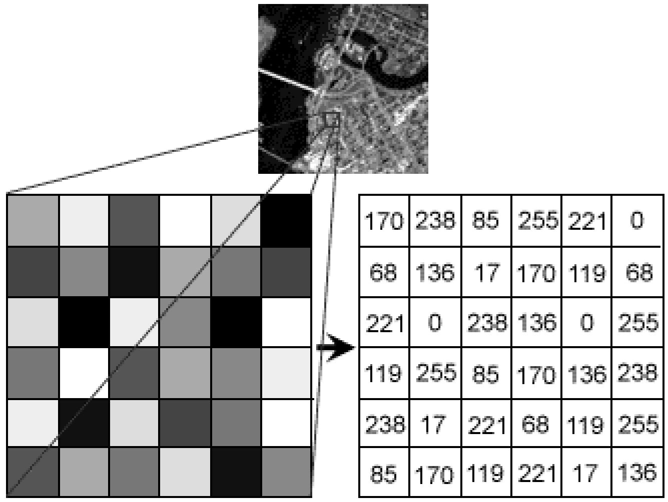

Fig 2: image in form of pixels

Each pixel has a color. The color is a 32-bit integer. The first eight bits determine the redness of the pixel, the next eight bits the greenness, the next eight bits the blueness, and the remaining eight bits the transparency of the pixel.

1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1
Transparency	Red	Green	Blue

Fig 3: Pixel colour

II. EXISTING WORK

Image Acquisition is to acquire a digital image. To do so requires an image sensor and the capability to digitize the signal produced by the sensor. The sensor could be monochrome or color TV camera that produces an entire image of the problem domain every 1/30 sec. the image sensor could also be line scan camera that produces a single image line at a time. In this case, the objects motion past the line.

Fig 4: line scan camera

Scanner produces a two-dimensional image. If the output of the camera or other imaging sensor is not in digital form, an analog to digital converter digitizes it. The nature of the sensor and the image it produces are determined by the application.

ISSN- 2394-5125 VOL 06, ISSUE 01, 2019

Fig 5: scanner

Image enhancement is among the simplest and most appealing areas of digital image processing. Basically, the idea behind enhancement techniques is to bring out detail that is obscured, or simply to highlight certain features of interesting an image. A familiar example of enhancement is when we increase the contrast of an image because "it looks better." It is important to keep in mind that enhancement is a very subjective area of image processing.

Fig 6: Enhancement of an image

Image restoration:
Image restoration is an area that also deals with improving the appearance of an image. However, unlike enhancement, which is subjective, image restoration is objective, in the sense that restoration techniques tend to be based on mathematical or probabilistic models of image degradation.

Enhancement, on the other hand, is based on human subjective preferences regarding what constitutes a "good" enhancement result. For example, contrast stretching is considered an enhancement technique because it is based primarily on the pleasing aspects it might present to the viewer, where as removal of image blur by applying a deblurring function is considered a restoration technique.
Color image processing:
The use of color in image processing is motivated by two principal factors. First, color is a powerful descriptor that often simplifies object identification and extraction from a scene. Second, humans can discern thousands of color shades and intensities, compared to about only two dozen shades of gray. This second factor is particularly important in manual image analysis.

Fig 7: Gray image and color image

ISSN- 2394-5125 VOL 06, ISSUE 01, 2019

Wavelets and multiresolution processing:

Wavelets are the formation for representing images in various degrees of resolution. Although the Fourier transform has been the mainstay of transform based image processing since the late1950's, a more recent transformation, called the wavelet transform, and is now making it even easier to compress, transmit, and analyze many images. Unlike the Fourier transform, whose basis functions are sinusoids, wavelet transforms are based on small values, called Wavelets, of varying frequency and limited duration.

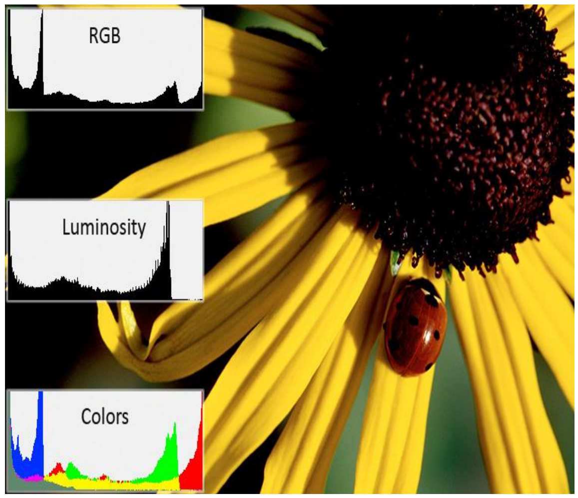

Fig 8: An image representing in RGB & luminosity form

Wavelets were first shown to be the foundation of a powerful new approach to signal processing and analysis called Multiresolution theory. Multiresolution theory incorporates and unifies techniques from a variety of disciplines, including sub band coding from signal processing, quadrature mirror filtering from digital speech recognition, and pyramidal image processing.

Compression:

Compression, as the name implies, deals with techniques for reducing the storage required saving an image, or the bandwidth required for transmitting it. Although storage technology has improved significantly over the past decade, the same cannot be said for transmission capacity. This is true particularly in uses of the Internet, which are characterized by significant pictorial content. Image compression is familiar to most users of computers in the form of image file extensions, such as the jpg file extension used in the JPEG (Joint Photographic Experts Group) image compression standard.

Morphological processing:

Morphological processing deals with tools for extracting image components that are useful in the representation and description of shape. The language of mathematical morphology is set theory. As such, morphology offers a unified and powerful approach to numerous image processing problems. Sets in mathematical morphology represent objects in an image. For example, the set of all black pixels in a binary image is a complete morphological description of the image.

Fig 9: Extracting of an image

In binary images, the sets in question are members of the 2-D integer space Z2, where each element of a set is a 2-D vector whose coordinates are the (x,y) coordinates of a black(or white) pixel in the image. Gray-scale digital images can be represented

ISSN- 2394-5125 VOL 06, ISSUE 01, 2019

as sets whose components are in Z3. In this case, two components of each element of the set refer to the coordinates of a pixel, and the third corresponds to its discrete gray-level value.

Segmentation:
Segmentation procedures partition an image into its constituent parts or objects. In general, autonomous segmentation is one of the most difficult tasks in digital image processing. A rugged segmentation procedure brings the process a long way toward successful solution of imaging problems that require objects to be identified individually.

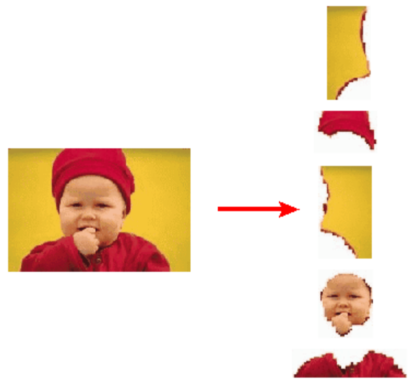

Fig 10: Segmenting of an image

On the other hand, weak or erratic segmentation algorithms almost always guarantee eventual failure. In general, the more accurate the segmentation, the more likely recognition is to succeed.
Representation and description:
Representation and description almost always follow the output of a segmentation stage, which usually is raw pixel data, constituting either the boundary of a region (i.e., the set of pixels separating one image region from another) or all the points in the region itself. In either case, converting the data to a form suitable for computer processing is necessary. The first decision that must be made is whether the data should be represented as a boundary or as a complete region. Boundary representation is appropriate when the focus is on external shape characteristics, such as corners and inflections.
Regional representation is appropriate when the focus is on internal properties, such as texture or skeletal shape. In some applications, these representations complement each other. Choosing a representation is only part of the solution for transforming raw data into a form suitable for subsequent computer processing. A method must also be specified for describing the data so that features of interest are highlighted. Description, also called feature selection, deals with extracting attributes that result in some quantitative information of interest or are basic for differentiating one class of objects from another.
Object recognition:
The last stage involves recognition and interpretation. Recognition is the process that assigns a label to an object based on the information provided by its descriptors. Interpretation involves assigning meaning to an ensemble of recognized objects.
Knowledgebase:
Knowledge about a problem domain is coded into image processing system in the form of a knowledge database. This knowledge may be as simple as detailing regions of an image when the information of interests is known to be located, thus limiting the search that has to be conducted in seeking that information. The knowledge base also can be quite complex, such as an inter related to list of all major possible defects in a materials inspection problem or an image data base containing high resolution satellite images of a region in connection with change deletion application. In addition to guiding the operation of each processing module, the knowledge base also controls the interaction between modules. The system must be endowed with the knowledge to recognize the significance of the location of the string with respect to other components of an address field. This knowledge glides not only the operation of each module, but it also aids in feedback operations between modules through the knowledge base. We implemented preprocessing techniques using MATLAB.

III. PROPOSED WORK

Background maintenance seems simple at first. But it turns out to be a problem rich with hard cases and subtle tradeoffs. This section attempts to crystallize the issues by
proposing a set of principles to which background maintenance modules should adhere. Principle 1: No Semantics On one hand, only Wallflower succeeds on the foreground aperture test, where movement of the interior of subject's shirt is undetectable. The region filling algorithm fills most of the area that the other algorithms miss. On the other hand, Wallflower outputs false positives in the waving trees sequence, where part of the sky, considered background by the pixel-level, becomes foreground after region-level processing. The region-level algorithm is therefore an unsound heuristic, the use of which is not justified in general because it is an attempt to extract object semantics from low-level vision:

Semantic differentiation of objects should not be handled by the background maintenance module. Background subtraction is never an end in itself. Larger systems seeking a high level understanding of image sequences use it as a component. Put another way, a background maintenance module handles the default model for everything in a scene that is not modeled explicitly by other processing modules. Thus, the module performing background maintenance should not attempt to extract the semantics of foreground objects on its own. This restriction implies that there are some problems that are inappropriate for background maintenance systems to attempt to solve. In particular, attempts to solve the sleeping person, waking person, and foreground aperture problems are bound to be unsound, given the low-level nature of background maintenance. A corollary to this principle is that while background maintenance might be useful in determining gross traffic statistics of objects such as people and cars, which can alternately be moving or motionless, attempts to use it alone as a preprocessing step for continuous, accurate tracking are bound to fail – some foreground objects will either be incorrectly adapted into the background or moved background objects will remain forever in the foreground. We note that in some systems, the higher-level module provides feedback to background maintenance about what pixels should not be adapted into the background. In one person-tracking system, a high-level module tracks a single user and prevents pixels representing the user from becoming background [10]. This is different from attempting to segment whole people, as such, in the background module itself; the background module merely accepts additional information from above but does not generate it on its own.

Principle 2: Proper Initial Segmentation

As long as there is to be a background maintenance module, we must differentiate the task of finding foreground objects from the task of understanding whether or not they are of interest to the system. Background subtraction should segment objects of interest when they first appear (or reappear) in a scene. Instead of semantic understanding, a more realistic goal for background subtraction is to pick out objects of interest that can then be recognized, tracked, or ignored by higher-level processing. In environments where adaptation is necessary (see Principle 4), maintaining foreground objects as foreground is not a reasonable task for background modelers, since such accurate maintenance requires semantic understanding of foreground. We can evaluate background maintenance by how closely it comes to finding all foreground pixels (as defined by the end task) when a foreground object first appears in the scene, while simultaneously ignoring all others. Object recognition and tracking can be computationally expensive tasks – good background subtraction eliminates the need to perform these tasks for each frame, on every sub region.

Principle 3: Stationarity Criteria

The principle of proper initial segmentation depends on the notion of object salience. How, in general, should one determine what is of interest in video sequences? Objects are salient when they deviate from some invariant property of the background. The key question, then, is how this invariance is modeled and what it means to deviate from it. Backgrounds, for instance, are not necessarily defined by absence of motion. Consider a fluttering leaf on a tree. As the leaf moves on and off a pixel, that pixel's value will change radically. No unimodal distribution of pixel values can adequately capture such a background, because these models implicitly assume that the background, apart from some minimal amount of noise, is static. Restrictions apply. unimodal models failed to capture the complexity in the background required to handle the waving trees and camouflage experiments. Those background systems which can model multimodal distributions of pixels, or which attempt to make predictions about the expected background fared well, however. These algorithms succeed because they are founded on a stronger notion of when a pixel deviates from the background. An appropriate pixel-level stationarity criterion should be defined. Pixels that satisfy this criterion are declared background and ignored. We define stationarity as that quality of the background that a particular model assumes to be approximately constant. (We note that this is different, and more ambiguous, than the formal definition of stationarity from random signal processing.) In order to understand the strengths and weaknesses of a particular algorithm in a specified environment, it is crucial that this stationarity criterion be made explicit. Carelessness in defining the stationarity criterion can lead to the waving trees and camouflage problems.

Principle 4: Adaptation

Backgrounds often change, even with liberal definitions of stationarity. For instance, all of the adaptive background maintenance systems are able to handle the moved object sequence by absorbing the chair into the background when it regains stationarity. Ergo, the background model must adapt to both sudden and gradual changes in the background. This is an obvious requirement, but it makes sense only under Principles 1 and 2. The line between gradual and non-gradual changes should be chosen to maximize the distinction between events that cause them. The surprisingly good performance of frame differencing suggests that fast adaptation works quite well if the foreground consists of moving people. Ultimately, however, the choice of this line is arbitrary. Some parts of a scene may remain in the foreground unnecessarily long if adaptation is slow, but other parts will disappear too rapidly into the background if adaptation is fast. Neither approach is inherently better than the other – a point that emphasizes the inadequacy of background maintenance for all but the initialization of tracking. This point is also consistent with Principle 2, which does not distinguish between algorithms that adapt quickly or slowly, as long as new foreground objects will still appear in the foreground.

Principle 5: Multiple Spatial Levels

Sudden light changes were best handled by normalized block correlation, Eigen background, and Wallflower. On the other hand, neither Eigen backgrounds nor block correlation deals with the moved object problem or the bootstrapping problem,

ISSN- 2394-5125 VOL 06, ISSUE 01, 2019

because they lack adaptive pixel level models. So, our final principle is the following: Background models should take into account changes at differing spatial scales. Most background maintenance algorithms maintain either pixel-wise models or whole-frame models, but not both. Pixel-level models are necessary to solve many of the most common background maintenance problems, while the light switch and time-of-day problems suggests that frame-wide models are useful, as well. A good background maintenance system is likely to explicitly model changes that happen at different spatial scales. Much of Wallflower's success is attributable to its separate models for pixels and frames.

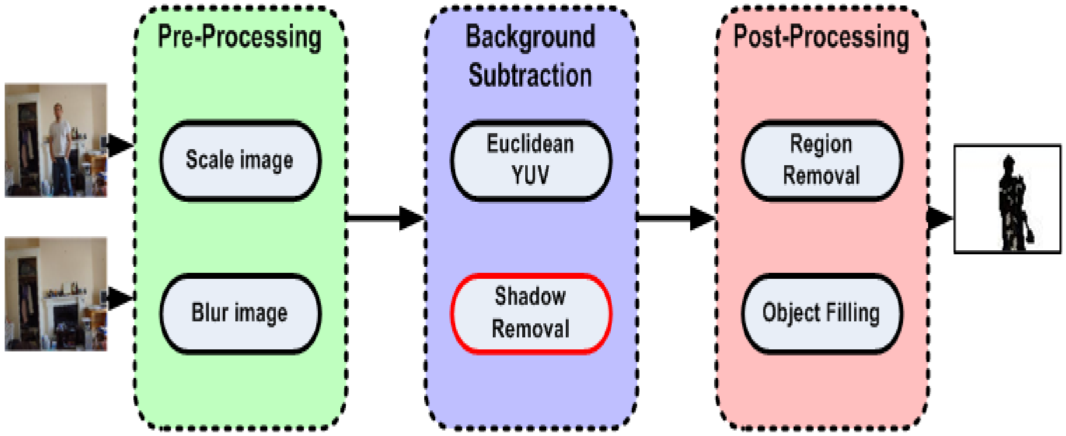

Fig 11: Segmenting of an image

IV.RESULTS AND DISCUSSION

HSL and HSV are simple transformations of device-dependent RGB, the color defined by a (h, s, l) or (h, s, v) triplet depends on the particular color of red, green, and blue "primaries" used. Each unique RGB device therefore has unique HSL and HSV spaces to accompany it. An (h, s, l) or (h, s, v) triplet can however become definite when it is tied to a particular RGB color space, such as RGB. The HSV model is commonly used in computer graphics applications. In various application contexts, a user must choose a color to be applied to a particular graphical element. When used in this way, the HSV color wheel is often used. In it, the hue is represented by a circular region; a separate triangular region may be used to represent saturation and value. Typically, the vertical axis of the triangle indicates saturation, while the horizontal axis corresponds to value. In this way, a color can be chosen by first picking the hue from the circular region, then selecting the desired saturation and value from the triangular region.

The conical representation of the HSV model is well-suited to visualizing the entire HSV color space in a single object.

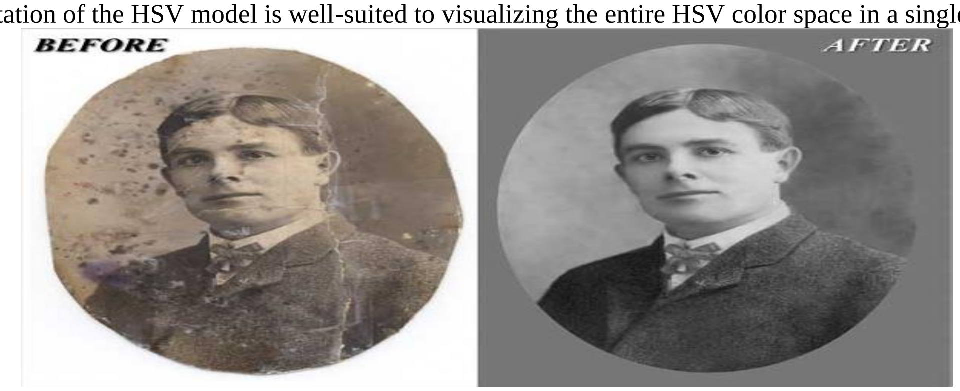

Fig 12: restoration of an image

V.CONCLUSION

The advantages of using color as feature to achieve object's similarity is analyzed and found that it is robust against the complex, deformed and changeable shape (i.e. different human profiles). In addition, it is also scale and rotation invariant, as well as faster in terms of processing time. Color information is extracted, stored and compared to find uniqueness of each object. Thus far, we have mainly deal with hardware and software setting up. After creating working environment, we have utilized some detection algorithms that extract moving objects from scenes with cluttered backgrounds. Next, we plan to write the necessary algorithms to realize object localization in the scene and artificial intelligence to track that moving object. Moreover, we will add a some kind of pattern matching algorithm to check if the detected object is a human face. Currently, the system works at

22-24 frames per second which can be considered to be real-time. We will try to keep up with that and optimize upcoming additions to the algorithm and preserve the real –time property of the system.

REFERENCES

[1] R. Cucchiara, C. Grana, G. Neri, M. Piccardi and A. Prati, “The Sakbot system for moving object detection and tracking,” Video-based Surveillance Systems-Computer vision and Distributed Processing, pp. 145-157, 2001.

[2] 2. C. Stauffer and W. E. L. Grimson, “Adaptive background mixture models for real-time tracking,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 1999.

[3] 3. C. Wren, A. Azarbayejani, T. Darrell, A. Pentl, “Pfinder: Real-time tracking of the human body,” In IEEE Trans. Pattern Analysis and Machine Intelligent, vol. 19, no. 7, pp. 780-785.

[4] 4. L. Qiu and L. Li, “Contour extraction of moving objects,” in Proc. IEEE Int’l Conf. Pattern Recognition, vol. 2, 1998, pp. 1427–1432.

[5] 5. Tang Sze Ling, Liang Kim Meng, Lim Mei Kuan, Zulaikha Kadim and Ahmed A. Baha‘a Al-Deen, “Colourbased Object Tracking in Surveillance Application” in Proceedings of the International MultiConference of Engineers and Computer Scientists 2009 Vol I IMECS 2009, March 18 - 20, 2009, Hong Kong.

[6] 6. S. J. McKenna, S. Jabri, Z. Duric, A. Rosenfeld and H. Wechsler, “Tracking group of people,” Comput. Vis. Image Understanding, vol. 80, no. 1, pp. 42-56, 2000.

[7] 7. I. Haritaoglu, D. Harwood, and L. S. Davis, “W4: Real time surveillance of people and their activities,” In IEEE Trans. Pattern Analysis and Machine Intelligent, vol. 22, no. 8, 2000, pp. 809-830.

[8] 8. A. Lipton, H. Fujiyoshi and R. Patil, “Moving target classification and tracking from real-time video,” In DARPA Image Understanding Workshop, pp. 129-136, November 1998.

[9] 9. P. Pérez, J. Vermaak, and A. Blake, "Data fusion for tracking with particles," Proceedings of the IEEE, vol. 92, no. 3, pp. 495-513, (2004).

[10] 10. C. Shen, A. van den Hengel, and A. Dick, "Probabilistic multiple cue integration for particle filter based tracking," in Proc. of the VIIth Digital Image Computing: Techniques and Applications. C. Sun, H. Talbot, S. Ourselin, T. Adriansen, Eds., 10-12, (2003).

[11] 11. Wang, H., et al., "Adaptive object tracking based on an effective appearance filter". IEEE Transactions on Pattern Analysis and Machine Intelligence,(2007).