

報告番号	※甲	第	号
------	----	---	---

主 論 文 の 要 旨

論文題目	Visual scene analysis based on the kurtosis of responses of independent component filters
氏 名	ANDRE BORGES CAVALCANTE

論 文 内 容 の 要 旨

Background: An old problem in information science and engineering is visual scene analysis. In visual scene analysis, the goal is to test some hypothesis or to conclude some fact about a visual scene by using image processing techniques. Generally, visual scene analysis is achieved by the implementation of two perceptual processes, i.e., segmentation and recognition. Segmentation consists of segregating the visual scene into different regions or partitions. Recognition consists of classifying the visual scene partitions or the entire visual scene into categories defined a priori.

Segmentation and recognition are the base of many applications such as surveillance, medical image processing and assisted diagnostic, biometrics, object detection and tracking, visual inspection, text and document handling, etc. These applications are implemented in different platforms such as automotive, airborne and space, hospital and medical equipment, and general low-power consumer electronics such as smartphones and tablets. These platforms involve different trade-offs between energy consumption, time consumption, memory consumption, size and weight of the processor device. Therefore, there is a need for methods which can fast perform visual scene analysis with low computation.

Goal: The goal of this thesis is to introduce a new methodology for visual scene analysis which has fast execution and low computational cost. Our methodology consists of analyzing the kurtosis of responses of independent component (IC) filters. This methodology can be used to implement both segmentation and recognition processes.

Methodology: In this thesis, we propose to use the kurtosis of responses of IC filters for visual scene analysis. In this way, our methodology is mainly based on two concepts, which are independent component filters and kurtosis. Independent component filters are learned so as to have mutually statistically independent responses. Kurtosis is defined as the standardized fourth-order moment of a random variable.

In the proposed methodology, we calculate the responses of two different sets of IC filters to the same input image. Ideally, the first set of filters is designed to respond strongly to only one type or category of image defined a priori. Similarly, the second set of filters is designed to respond strongly to a type of image different than that of the first set of filters. By analyzing the kurtosis of all filter responses, our methodology is able to segment or recognize different type of regions in the input image.

Methods and Experiments: Our methodology is applied on three problems or subjects of visual scene analysis. The first subject is segmentation of depth-of-field images. The second subject is segmentation of nature-made and man-made structures. The third subject is measuring the perception of complexity in streetscapes.

In the first studied subject of visual scene analysis, i.e., segmentation of depth-of-field images (Chapter 3), the goal is to recognize and segment focused and unfocused regions in the visual scene. In our proposed method, two different sets of IC filters are united or joined together. One set of IC filters is designed to respond to focused image regions. The other set of IC filters is designed to respond to unfocused image regions. By measuring the kurtosis of responses of the whole filter population, our method is able to recognize which region in the scene is focused or unfocused. Here, we evaluate the performance of the proposed method for two different databases of depth-of-field images. And we use the classic F-measure as an objective criterion of performance. In this regard, our method exhibits the highest performance among the fast methods of segmentation of depth of field in terms of the F-measure. The performance of our method is only lower in comparison to that of time-consuming methods.

In the second studied subject of visual scene analysis, i.e., segmentation of natural and man-made structures (Chapter 4), the goal is to recognize and segment natural and artificial objects in the visual scene. Similar to the method of depth-of-field segmentation, two different sets of IC filters are united or joined together. One set of IC filters is designed to respond to natural objects. The other set of IC filters is designed to respond to artificial or man-made objects. By measuring the kurtosis of responses of the whole filter population, our method is able to recognize whether each region in the image is natural or artificial. We evaluate the performance of the proposed method for two different types of datasets. The first dataset consists of images acquired by still-devices. The second dataset consists of images acquired by moving devices. Each dataset consists of two different image databases. Here, we also use the F-measure as an objective criterion of performance. In comparison to other methods, our proposed system exhibits the highest performance in terms of the F-measure.

In the third studied subject of visual scene analysis, i.e., measuring the perception of complexity in streetscapes (Chapter 5), the goal is to create a measure to quantify the

human perception of visual complexity in streetscapes. Our proposed measure of complexity is based on the statistics of local contrast and spatial frequency of the visual scene. Notice that the kurtosis-based methodology proposed in this thesis is used to extract the statistics of local spatial frequency. For evaluation of performance of our method, we have built a dataset of streetscape images. This dataset consists of images acquired in four cities. Two cities in Algeria and two cities in Japan.

For evaluation of performance, we use the correlation coefficient between complexity ratings from human participants and the objective measures of complexity. In comparison to classic and new methods, the proposed measure of complexity exhibits higher correlation with the subjective rating of human participants. Furthermore, we show that our measure can be used to analyze perception in nighttime images.

Conclusion: This thesis introduced a new methodology for visual scene analysis. This new methodology is general and it can be applied in many subjects or problems of visual scene analysis. Based on the proposed methodology, new computer systems are created for applications such as segmentation of depth-of-field image, segmentation of natural and man-made structures, and measuring the perception of complexity in streetscapes. For all these three subjects, this thesis presents motivation, the state of the art, open problems, and finally our experiments and results using the proposed methodology. In this regard, our systems exhibit competitive or higher performance than that of the state of the art for all subjects. It is also important to highlight that the proposed methodology is fast and low computational costly. Therefore, our methodology is attractive for industry and consumer real-time vision applications.

Future work can include the extension or modification of the methodology by using a hierarchical architecture, i.e., architecture with several layers of filtering. Furthermore, future work include applying the proposed systems for problems related to driving and navigation.

