Real-time Image Recovery Using Temporal Image Fusion

André Mora, José Manuel Fonseca, Rita Ribeiro
Center of Technology and Systems, Uninova, Monte da Caparica, Portugal
Dept. of Electrotechnical Engineering, Faculty of Sciences and Technology, UNL, Monte da Caparica, Portugal
atm@uninova.pt, jmf@uninova.pt, rar@uninova.pt

Abstract— In computer vision systems an unpredictable image corruption can have significant impact on its usability. Image recovery methods for partial image damage, in particular in moving scenarios, can be crucial for recovering corrupted images. In these situations, image fusion techniques can be successfully applied to congregate information taken at different instants and from different points-of-view to recover damaged parts. In this article we propose a technique for temporal and spatial image fusion, based on fuzzy confidence classification, which allows partial image recovery upon unexpected defects without user intervention. The method uses image alignment techniques and duplicated information from previous images to create fuzzy confidence maps. These maps are then used to detect damaged pixels and recover them using information from previous frames.

Keywords—Image fusion, fuzzy confidence, spatial-temporal fusion, image registration

I. OVERVIEW

In image processing scenarios where real-time information extraction is required and unpredictable image corruption can occur, image fusion techniques can be of significant importance, especially for unmanned vehicles guidance, such as, aerial vehicles [1], [2], mobile robots [3], [4] or even for planetary landing [5–7]. For example, in the image corruption scenario, shown in Figure 2, image fusion techniques can be applied by gathering information from the previous instants to recover information for current damaged pixels. Image corruption may be caused by three main factors: originated by dust in the image sensor causing sharp dark spots over the image in fixed positions; dirty lens causing blurred regions on the image; or light reflections inside the lens (also known by light flares), which produce light rounded spots that can be moving on the resulting image.

The main objective of image fusion is to reduce uncertainty and redundancy while maximizing relevant information, by combining different image representations of the same scene [8]. The general procedure is to use several images of the same scene, provided by different sensors, and combine them to obtain a complete view of the scene, not only in terms of position and geometry, but also in terms of semantic interpretation. In this work, it is proposed a spatial-temporal image fusion architecture with the aim of recovering corrupted images.

Image fusion algorithms can be divided into pixel, feature, and symbolic levels. However, feature and symbolic algorithms have not received the same level of attention as pixel-level algorithms. Pixel-level algorithms are the more common and they work either in the spatial domain or in the transform domain. Although pixel-level fusion is a local operation, transform domain algorithms create the fused images globally. Feature-based algorithms (symbolic) typically segment the images into regions and fuse the regions using the images intrinsic properties [9], [10]. This study is focused on pixel-level methods and techniques.

A general architecture for a spatial-temporal image fusion technique is discussed in this paper, which gathers information along several consecutive images to produce pixel-level confidence maps and then enabling fixing damaged pixels. The technique is divided into three main steps: image registration process; temporal fuzzy confidence map generation; and image recovery process.

II. IMAGE FUSION ARCHITECTURE

The image fusion architecture proposed can be considered as a spatial-temporal image fusion technique, since it gathers information along several consecutive images to produce pixel-level confidence maps and, afterwards, fix damaged pixels. The diagram in Figure 1 shows the overall architecture. It is divided into three main steps:

The first step is image alignment which is responsible by adjusting zoom, rotation and panning of the new image using image registration techniques or using navigation information (altitude, gyros, etc.) that provides the current camera viewpoint. The objective is to have the last k images aligned with the current image.

The second step is the fuzzy confidence map generation. The inputs are the registered image, the previous images and their respective confidence maps. These inputs will be analyzed to detect damaged pixels and producing as output the current confidence map. The pixel’s confidence will reflect qualitatively the probability of being a damaged pixel.

Figure 1. Image spatial-temporal fusion architecture.
Finally, in the last step the pixels from the registered image with low confidence are fused with information from previous images. The inputs are the previous and current images and their corresponding confidence maps and the output is the fused recovered image.

III. IMAGE REGISTRATION

For the image alignment, a pre-processing step involving image registration is required to achieve spatial correspondence between the images. This image registration process is a crucial step to evaluate the geometrical transformations needed to align two or more images. Four groups of image registration techniques have been identified, depending on the image acquisition process [11]: multiview analysis (images taken from different viewpoints), multitemporal analysis (images taken at different times), multimodal analysis (images acquired by different sensors) and scene to model registration (the image is aligned to a previous model). The multitemporal analysis is the one applicable to our scenario. For a better alignment, the camera distortion model should also be available for the geometrical projection of previous images on the current field-of-view.

IV. FUZZY CONFIDENCE

The detection of incoherent situations by pixel confidence level evaluation is an important feature of any autonomous computer vision system. Fuzzy logic techniques [12], [13] are well suited for measuring qualitatively this confidence level since it can be derived from its pixel intensity, its neighbors’ intensities and/or from its previous confidence levels. For its computation temporal fusion, using previous images, will be taken into consideration (Figure 1).

The algorithm parameters are defined with fuzzy membership functions to identify what is a low, medium or high confidence level. These functions can be obtained either empirically or by using any learning techniques such as neural networks [10], [12].

Predicting the next image $t_k$ based on Kalman filters is also envisaged to be studied in order to improve the fuzzy confidence classifier. However, the trade-off between the advantages of using more complex techniques and the computational overhead load should be considered.

The inputs to calculate the fuzzy confidence $C_t(x,y)$ for the pixel $I_t(x,y)$ are:

- $I_t(x,y)$ intensity
- $I_{t-1}(x,y) .. I_{t-n}(x,y)$ intensity and correspondent confidence $C_{t-1}(x,y) .. C_{t-n}(x,y)$ levels
- $I_t(x,y)$ neighbors average intensity
- $I_{t+1}(x,y)$ predicted intensity

These inputs represent both temporal (by using the previous and the next image) and spatial information (through the use of the neighbors’ intensity) used. Spatial information is important to achieve better noise independence [13].

In Figure 3 we present an example of the noise detection algorithm and corresponding pixel confidence level generation. As it can be seen, the system will use a set of archived images (a1 to a3) that in each cycle will be projected to the coordinates of the newly acquired image, obtaining the aligned images (b1 to b3) that should match each other.

However, when significant noise is present, images are substantially different from their projected counterparts, creating suspicion about some areas (c1 to c3). The evaluation of the different images will be consolidated on a single difference image (d1) that is the basis for the creation of the pixel confidence levels (d2). In the pixel confidence image (Figure 3.d2) the colors white, yellow and red represent respectively high, medium and low confidence pixels.
V. IMAGE RECOVERY: ILLUSTRATIVE EXAMPLE

The image corruption presented can produce dark spots on the image, thus generating shadows inexistent on the real scenario (Figure 2). In a moving application such as UAV landing one, these shadows will remain fixed within the image layout, but the real view of the terrain can be updated as it gets out from the dark areas.

As shown in the proposed architecture (Figure 1) in the recovery process a small subset of historic images (typically 3 to 4 images and confidence levels) will be combined (fusion) with the current one into a “recovered”, improved image. Having access to the past images and their respective confidence information, provide the means for recovering low confidence regions using either classical interpolation techniques or novel approaches such as specialized aggregation operators [14].

Now we illustrate a fusion recovery process using a simple interpolation technique. For the recovery algorithm a small subset of images, typically 3 to 4 images, will be stored in memory among their correspondent confidence levels, mainly for determining confidence levels. In Figure 4, the presence of low confidence levels (b2) on the newly acquired image (b1) motivates a recovery algorithm that begins by projecting the archived image (a1), that also has low confidence pixels (a2) to the coordinates of the new image resulting on image (c1) and corresponding confidence levels (c2). The confidence levels of (a1) are adjusted (in this case they are downgraded from high to medium) because the image is interpolated to produce an image with the same size as (b1) with the same transformation applied to its confidence levels.

Figure 3. Pixel confidence levels generation; a1 to a3) Image obtained in t-2, t-1 and current cycle; b1) image equivalent to a1 obtained from the projection of the image t-3; b2) idem with a2 and t-3; b3) idem with a3 and t-3; c1 to c3) difference between a1 and b1, a2 and b2, a3 and b3; d1) total images differences (sum of c1 to c3) d2) final pixel confidence levels (white is high confidence, yellow medium confidence and red low confidence).
The two images are compared and low confidence pixels of (b1) are replaced by the corresponding pixels on (c1) if they have a medium or high confidence level. The recovered image is presented on (d1) with its confidence levels on (d2). As it can be seen, a reconstructed image can be obtained by this process with automatic detection of noisy areas.

It must be stressed that the level of reconstruction will depend on the type of noise, the performance of the alignment between the images and the number of archived images considered for the reconstruction process. In this example the recovered image (d1) keeps some noisy areas because they coincide on both the captured and projected images and the confidence levels (d2) show a limited trust on the reconstructed pixels.

Other approaches could be included as recovery process (step 3). An interesting novel approach for fusing images and information is based on upward and downward reinforcement operators [14]. These specialized operators have already proven to be effective for data fusion of hazard maps in spacecraft safe landing [14], [15] and they will be considered in future developments of this work.

Finally, the reconstructed/fused image together with its pixel confidence levels, are then ready to be passed to the following stages of any computer vision system.

VI. CONCLUSIONS

In this paper, we presented an architecture for temporal image fusion, which allows the detection and correction of faulty pixels/regions. In some real-time applications repairing faulty image acquisition areas cannot be done due to the difficulty in getting to the equipment or because it is in critical operation. Therefore, real-time fusion processes that can detect and correct or overcome such defects are important.

The detection of faulty pixels is not a trivial operation. It requires the evaluation of several conditions, such as the pixel value should not change during several iterations, but it also can be due to noise. Therefore, we discussed a novel image fusion architecture, which includes a fuzzy decision classifier to evaluate pixels’ confidence level.

The recovery step (fusion of images) will calculate the confidence level from the current and the previous images. The low confidence pixels on each image are replaced by pixels in the same position on the previous images whose confidence was high. Using this technique damaged pixels can be recovered providing an improved image to the computer vision system, hopefully free of damaged pixels.

In future works we will consider image fusion using specialized reinforcement aggregating operators for the recovery process, as discussed in section V, because they were successfully applied in spacecraft safe landing site selection with hazard avoidance.
REFERENCES


