Fusing images for contextual enhancement

Shahan C. Nercessian, Karen Panetta, and Sos Agaian

A new image fusion algorithm that uses human visual characteristics and a novel adaptive weighting scheme can produce more visually relevant image combinations from different modalities.

Imaging sensors have different capabilities, strengths, and weaknesses, so it is natural to consider combining their advantages into a single, fused image representation. The aim of image fusion algorithms is to combine the similar and dissimilar information of images obtained using different capture techniques, thus improving the perception of the observed object or scene. Image fusion already has aided medical diagnostics and improved visibility for military, security, and remote sensing applications.

Multi-scale transform domain image fusion algorithms, such as the wavelet and pyramidal transform-based approaches, are of particular interest because they provide a way to merge the multi-scale edge structures of the source images. In those approaches, one fusion rule melds coarse-level information from the respective source images, while a second fusion rule combines multi-scale edge information. The former has been achieved by either a global averaging rule—which is generally not local enough—or by an adaptive means, which to date has introduced artifacts into the fused result. Multi-scale edge structures have generally been combined using an edge definition that does not account for various human visual system (HVS) phenomena.

It is harder to fuse images with dissimilar information, because it is more difficult to determine which image should be more heavily weighted at a given location. However, this is precisely when image fusion may provide its greatest benefit. We have been working on a new algorithm that uses a similarity-based adaptive weighting scheme to fuse coarse-level information. This localizes the weighting, and no artifacts are introduced into the fusion result. Also, leveraging HVS as an image processor, we fused multi-scale structures using a parametric, perceptually-driven contrast measure.

The coarse-level fusion rule’s goal is to maintain appropriate local luminance levels in the fusion result based on the source images’ local luminance. We developed a new adaptive fusion rule based on our observation that the source image containing more relevant information locally is generally apparent from the level-coarse information’s uniform average. Therefore, weights can be determined based on which source image looks more locally similar to this initial estimate, as determined by an objective image similarity metric. Figure 1 shows the benefits of our proposed coarse-level fusion rule when applied directly in the spatial domain. The rule can adaptively weigh the fusion so the relevant structures are appropriately represented, without introducing artifacts in the processed result.

The aim of the detail fusion rule is to appropriately inject image contrast and edge structures. To this end, we used a multi-scale contrast measure that accounts for the HVS’ perceived contrast being a function of relative and not absolute contrast changes. The parametric nature of the contrast definitions allows

Figure 1. (a, b) Original navigation source images. Spatial image fusion results using (c) uniform averaging, (d) adaptive principal component analysis, and (e) presented method.

Continued on next page
the algorithm to tune the degree to which the masking effect holds.

Figure 2 shows a comparison between our multi-scale method and conventional spatial and transform domain algorithms. Our technique is able to provide more visually appealing fusion results. Quantitatively, our method yields image fusion results that are more relevant to the input images themselves. The presented coarse-fusion rule is able to appropriately determine the local luminance levels of the fusion, while the detail fusion rule allows image contrast to be combined in a perceptually-driven manner.

In conclusion, we have introduced a new, adaptive image fusion algorithm that is viable for both visually similar and dissimilar source images. We are particularly interested in using our fusion technology to combine information from different medical imaging modalities, and ultimately, to incorporate the tool into automated vision systems.

Author Information

Shahan C. Nercessian and Karen Panetta
Tufts University
Medford, MA

Sos Agaian
University of Texas at San Antonio
San Antonio, TX

References