Scene context dependency of pattern constancy of time series imagery

Glenn Woodella, Daniel J. Jobsona, Zia-ur Rahamb
andaNASA Langley Research Center, Hampton, VA
bOld Dominion University, Norfolk, VA

ABSTRACT

A fundamental element of future generic pattern recognition technology is the ability to extract similar patterns for the same scene despite wide ranging extraneous variables, including lighting, turbidity, sensor exposure variations, and signal noise. In the process of demonstrating pattern constancy of this kind for retinex/visual servo (RVS) image enhancement processing, we found that the pattern constancy performance depended somewhat on scene content. Most notably, the scene topography and, in particular, the scale and extent of the topography in an image, affects the pattern constancy the most. This paper will explore these effects in more depth and present experimental data from several time series tests. These results further quantify the impact of topography on pattern constancy. Despite this residual inconstancy, the results of overall pattern constancy testing support the idea that RVS image processing can be a universal front-end for generic visual pattern recognition. While the effects on pattern constancy were significant, the RVS processing still does achieve a high degree of pattern constancy over a wide spectrum of scene content diversity, and wide ranging extraneousness variations in lighting, turbidity, and sensor exposure.

Keywords: Retinex, image enhancement, pattern recognition, edge detection

1. INTRODUCTION

NASA’s 2006 Strategic Plan, Sub-goal1, calls for the development of technologies for safer aircraft and higher capacity airspace systems. The trend is to get more aircraft down on the ground and back up in the air again as quickly as possible as existing airports are increasingly taxed to capacity. Of course increasing the density of the airspace has a potential negative affect on the safety of the systems. NASA’s Aviation Safety Program is tasked with making sure that safety is not compromised as capacity increases.

The advance of technology and the pressure to improve system capacity are bringing about change in cockpit systems. As new capabilities, information, and features are added, there is one constant, the pilot, who remains unchanged and must take on ever-increasing tasks and responsibilities. Pushed to the limit, the pilot may make mistakes or miss important details, placing undue risk on the aircraft and its passengers. Aviation accidents have been caused by pilot overload and may have been prevented had even one detail not been been missed or misunderstood. For this reason, systems and functions must be adapted to the pilot’s capabilities and allow for any limitations. The flight decks of the future must sense internal and external hazards, evaluate them, and provide key information to facilitate timely and appropriate responses by the pilot. This “smart” flight deck approach is one of NASA’s tools for creating safer skies. Called the Integrated Intelligent Flight Deck (IIFD), it provides a) advanced crew/vehicle interface technologies, b) technologies to enable detection of unsafe behaviors, c) fail-safe methods for changing the operator/automation roles in the presence of detected disability states, and d) a comprehensive surveillance system for the detection of external hazards with sufficient time-to-alarm for safe maneuvering to avoid the hazards. Our work in the Visual Information Processing Lab is concerned with this last element.

External hazards fall into one or more of the following categories: (1) meteorological (e.g. icing conditions, convective weather, wind gusts, turbulence), (2) environmental (e.g. volcanic ash), (3) geospatial (e.g. terrain, man-made obstacles, foreign object debris), (4) traffic, and (5) airspace constraints/restrictions. Imaging sensors can be effective against many of these hazards and pattern recognition is expected to play a crucial role in advanced, automated imaging systems. Whether it be for the detection of terrain features in the path of the aircraft, hazards on or around a runway, or other aircraft in flight, the automatic detection of such hazards is the logical next step in improving
the safety of our airways.

Effective and consistent pattern recognition requires that we be able to extract similar patterns for the same scene despite wide ranging extraneous variables, including lighting, turbidity, sensor exposure variations, and signal noise. While performing a variety of tests on various scenes in varying conditions we clearly noticed that the ability to match a scene under different conditions varied greatly depending the content of the scene. For this reason we decided to try to characterize this dependency of pattern constancy on scene context.

Both enhancement and canonical representation performance as well as pattern constancy performance are quite good over a wide range of extraneous imaging variabilities but at the extreme, are limited by sensor noise (or lossy image coding artifacts if present). Lesser limits on constancy and other performances are variable shadow edges, loss of small detail in small shadows, and any losses due to signal saturation and clipping (such information is irretrievably lost). Without image enhancement, there would be marginal or no correlation. For this reason, pattern constancy is demonstrated to a significant degree down to the fundamental limit of sensor noise and the usually lesser scene context dependencies explored here.

Topography is defined for this work as local relief in the scene which produces illumination variations due to localized lighting variations. This is as opposed to a reflectance-only scene in which the lightness values are not due to local shadow effects. It has been shown that image enhancement greatly reduces the effect of these lighting variations due to scene topography. For this work, the Langley-developed Retinex with Visual Servo (RVS)²,³,⁴,⁵ was used to reduce the effects of topographically-induced lighting variations. The RVS processing is described, and its performance as a precursor to edge detection demonstrated as an automatic, powerful, and generic image enhancement method in previous publications among many others in our collection of publications at http://dragon.larc.nasa.gov/retinex/background/retpubs.html. Pattern constancy, in support of pattern recognition for the detection of aviation external hazards, has revealed that although enhancement did improve images considerably, most notably reducing the effects of lighting variations, (and thereby approaching a canonical visual representation of the arbitrary scenes), shadows still affected the ability to score high correlations for scene recognition. The trend was that increasingly small scale topography produced lower correlation coefficients while scenes with large scale topography generally showed little to no effect. For this reason, changing the scale of a scene, zooming in for example, affected the ability to maintain a good match since what was once large scale topography was now small scale and changed the detectable edges to a great extent. This brought about the need to try to characterize the extent to which the scale of a scene had on the detection of edges and to try to understand how scene topography affects the edge detection results and subsequent pattern recognition correlation coefficient values in time series studies.

2. FROM SCENE TO EDGE PATTERN

All imagery for this study was acquired and processed in a similar manner, the end goal being to take the images through to edge patterns to generate correlation coefficients for a quantitative assessment of pattern matching. Images were acquired by a tripod-mounted Canon EOS-1Ds Mark II digital camera, at full resolution of 16.6 megapixels, and saved initially in high quality JPEG image format and later converted to TIFF. Lenses used were the Canon EF 70-200mm f/2.8L IS USM, and Canon EF 50mm f/1.8. Series of images were taken of each scene under sun illumination, approximately every 30 minutes, as sun angle changed, generating topographically-induced lighting variations. Original images throughout the time series were registered in Adobe Photoshop CS2 and then cropped selectively to 510 by 510 pixels to include areas of particular lighting, scaling, and/or topography variations.

Imagery was processed with RVS to enhance contrast prior to edge detection and compress the dynamic range in a non-linear manner. Figure 1 shows original imagery carried through the RVS processing and then to edge patterns. Figure 2 shows the same edge patterns both with and without prior RVS processing. The RVS has a proven track record for producing a canonical visual representation of arbitrary scenes and arbitrary imaging and sensor conditions - changing lighting and atmospheric conditions such as turbid imaging, and sensor exposure variations.

A Langley-developed, zero-crossing two-scale edge detection method⁹,¹⁰ was applied to the enhancements in order to isolate meaningful edges. This filter operates on the image at twice the spatial density, providing one-half pixel
location accuracy. This adaptive, two-scale operator mimics both the perceptual level of human visual perception as well as the anatomical/electrophysiological performance of human vision\textsuperscript{11,12,13}. This couples a high acuity/low contrast sensitivity with a lower acuity/ high contrast sensitivity.

![Figure 1](image1.png)

**Figure 1.** Representative scenes, processed images, and edge patterns for time series imagery at t\textsubscript{0} and t\textsubscript{n}.

![Figure 2](image2.png)

**Figure 2.** Comparisons of edge patterns for time series imagery with and without RVS processing.

The correlation coefficient is defined in Equation 1 for MxN images. It is calculated between all the edge images in each time series, as shown later in Figure 11, gives us quantitative insight into the performance of the enhancement processing and its ability to boost the likelihood of a valid pattern recognition event. The basic process which these data describe is that a good edge representation of a scene is stored in memory and the attempt is made to correlate this edge image with others coming into the processing, from real-time sensor data for example.

The correlation coefficient is defined in Equation 1 for MxN images.

$$\rho_{i,2} = \frac{1}{M_1 M_2} \sum_{m_1=0}^{M_1-1} \sum_{m_2=0}^{M_2-1} (g_1[m_1, m_2] - \bar{g}_1)(g_2[m_1, m_2] - \bar{g}_2) / \sigma_1 \sigma_2$$

From previous experiments with increasing injections of noise in test images, we found that the noise limit for correlation is \(r=0.00-0.04\) and for non-matching patterns range from \(r=0.00\) to \(r=0.05\). This supports the idea that a significant correlation value can be as low as \(~0.1\) or even less, perhaps down to 0.06.

**3. THE TEST SCENES**

Several different test scenes were used which ranged from featureless, with no variations in lighting and shadow
coverage to those with heavy topography and large variations in lighting and shadow coverage, to encompass a wide range of scales and types (i.e. high relief and more “rolling” smoother) of topography).

Scene one was the side of a large storage tank for liquid nitrogen, used as an operating atmosphere for a wind tunnel, and is shown in Figure 3. This scene has very little topography with the exception of a staircase which has a regular pattern casting a shadow across the curved tank surface. Four different series were studied. Two in which the exposure changed greatly over a relatively short period of time due to localized cloud cover and two where the exposures were fairly constant but the shadows cast across the surface varied greatly with time. In each of these, a smaller subset was taken of the larger to vary the scale.

![Figure 3. Test scenes of liquid nitrogen tank. The boxes indicate the region from where the subsets were taken.](image)

The second scene, shown in Figure 4, was the base of a tree which had very rough and textured bark. As the sun angle changed, the shadows moved across the scene, changing the lighting drastically. Two series were studied. One being a subset of the other to vary the scale as shown by the white box.

![Figure 4. Test scene of tree bark showing location where data subset was taken](image)

The third scene, shown in Figure 5, was the base of a light pole which had some moderate topography and was taken with changing sun angle. No subset of this scene was studied since it already contained small-scale topography at this scale.
Scene four, shown in Figure 6, was of the face of stone aggregate steps under varying sun angle. The minimal relief of the small stones provided a subset with limited amount of shadow-inducing topography and a more global scene with almost no shadows.

The fifth image set, shown in Figure 7, took advantage of a scene with very mixed topography. The global series of this brick wall contained a patch of rough brick mixed in with smooth brick. The subset consisted of only the rough brick.

The sixth series, shown in Figure 8, consisted of a scene under varying sun angle but with very little coarse topography. Being the side of a trailer, most of the image variation is reflectance-only with very few features. For this reason, no subset was taken.
The final scene, shown in Figure 9, again provided two sets. One of a smooth brick wall taken from from a distance so as to provide limited shadow effects between the bricks as the sun angle changed and the subset which had almost no topography-induced effects.

4. RESULTS

The first image subset was a series of six images taken over a relatively short period of time so as to minimize sun angle change but to capture a variety of exposures varying from strong sun to totally obscured sun. The results are illustrated in Figure 10.
Six images from each subset were used from the spatially registered set. Each image was processed with the RVS and then edge detection was performed. From the resulting edge images, correlation coefficients were calculated. Although the correlations were calculated on the edge pattern images, only the original images are shown in the chart since high correlations would naturally be the result of edge patterns which looked almost identical. By displaying the original images along with the correlation coefficients, one can get a better idea of how robust the RVS processing is over varying conditions.

To illustrate the utility of the edge pattern for scene recognition, we present a different scene. Figure 11 shows a set of correlation coefficients taken from edge patterns and compared with correlation coefficients taken from the original, unenhanced scene images. Correlation of imagery in the lightness domain yields higher values than does edge patterns but these higher values are less meaningful since edge patterns being binary are affected so much more by pattern mismatch, noise, and lighting variations. Correlation in the lightness domain yielded an average of 0.77 with a high of 0.96 and a low of 0.58. This high is very misleading and is supported by calculations on other examples which show values of 0.68 and greater for entirely different scenes.

For the nitrogen tank scene, edge correlation values ranged from a low of 0.32 to a high of 0.68 with all of the edge patterns showing very good correlation, indicating a good scene match. The mean correlation coefficient for the set was 0.45. The subset series was about half the size overall of the full set. Again, it showed a relatively high degree of correlation with the same mean of 0.45 with a high of 0.68 and a low of 0.31.

A second subset of the nitrogen tank series was compared. For this series, the exposures were nearly constant but the entire set was captured over a longer period of time so as to consider the shadow of the stairs that propagated across the tank throughout the morning. Again, the mean remained the same at 0.45 for the a series and 0.48 for the zoomed-in series. Highs were 0.58 and 0.60 and the lows were 0.33 and 0.42 respectively.

With the RVS processing taking care of exposure variations we fully expected the exposure series to score highly. but we expected the shadow series to show some degradation in the correlation as the heavy shadow traversed the tank. But this was not the case. We decided to further “degrade” the scene by reshooting the series but doing so from a slightly different point of view so as to get a truly different scene for comparison. The results are shown in Figure 12.
The results of this test were surprising but informative. We compared three images taken from three slightly different locations against one from the original scene. The correlation coefficients were again surprisingly high even though they were greatly diminished. While we were expecting values far less than what we had thought of as a lower scene detection threshold of 0.06, we actually got an average value of 0.08 which would indicate a scene match. When we went back to the edge patterns and examined them we realized that due to the regular pattern of the stairs, we were getting numerous false matches. Although edges were indeed lining up, they were not the same edges, but others that happened to coincide due to the many regular lines in the pattern. Figure 13 shows gray edge patterns from two different (but very similar scenes) overlaid. The white lines indicate correlating edges, many of which are not true matches but which add to the high correlation coefficient for this scene.

The scene with the greatest topography was that of the lower portion of a tree with very textural bark. Under differing sun angles, the changing shadows provided almost completely different scenes for comparison. Correlation values ranged from 0.16 to 0.52 with an average of 0.32 for the series and from 0.05 to 0.26 for an average of 0.14 for the zoomed-in series. Although one coefficient fell below our defined scene-recognition limit, the others were well above this threshold suggesting a scene match. As expected, this was the lowest scoring scene due to the amount of heavy topography on both the macro and micro scale.

The pole scene had plenty of small scale topography even at the macro scale so we only used one view. It had a low correlation coefficient of 0.16 and a high of 0.53 for an average of 0.31. The topography of this scene consisted of very definite edges which cast strong shadows across the smooth lower surface. The aggregate steps scored quite well on the macro scale but dropped significantly on the micro scale. As expected, the macro average was 0.43 while the subset scored as average of 0.27. Values ranged from 0.16 to 0.74 and 0.11 to 0.77 respectively. There were several high values, probably due to the almost regular pattern of the small stones, supporting the idea that heavy texture or regular patterns reduce apparent scene uniqueness. The white brick scene consisted of a large, smooth surface with a smaller, textural one. Between the bricks, a deep crevice provided some changing shadow detail. The zoomed-in subset included only the textural region and as such, scored below that of the larger scene. Values ranged from 0.17 to 0.66 with an average of 0.34 for the combined texture scene and 0.05 to 0.42 with an average of 0.21 for the coarse texture-only subset.
We performed a sub-test of this scene to assess the best case for our technique. The results are shown in Figure 14. Three images were captured at times $t_0$, $t_1$, and $t_2$ where $t_1 = t_0 + 3\text{sec}$ and $t_2 = t_1 + 60\text{sec}$.

![Figure 14. Correlation coefficients for short duration time lapse. Original images only shown for clarity.](image)

Figure 14 shows two of the short duration test images and the resulting difference image made by subtracting their edge images. The shadow has moved enough, even in a relatively short period of time, to lower the correlation significantly.

![Figure 15. Edge pattern subtraction from correlation limit test. Original images shown to illustrate near-identical scenes.](image)

Since we wanted a scene with very little small-scale topography, dominated by reflectance-only differences, we chose the side of a storage trailer which had very little real texture. As expected, the average was quite high, at 0.51 with a range of 0.32 to 0.69. Hardware attached to the trailer as well as rivet heads on the side wall, cast strong shadows across the scene as the sun illuminated the surface from a low angle which reduced the correlation coefficient significantly. These edges can easily be seen at the bottom of the right image of Figure 13. Rivet heads can be seen running vertically along the left edge. The area above the text consisted of white paint that had flaked off, leaving the darker, silver-gray aluminum exposed.

![Figure 16. Edge patterns showing small, strong shadows on a scene dominated by reflectance rather than by topography.](image)

The last scene consisted of the brick wall of a building taken from a far enough distance so as to lose fine shadow detail completely. This provided us with a macro scene in which a large shadow traversed across the scene as sun angle changed and a zoomed in subset where no shadows were involved. Both scenes showed very high correlations, with the macro scene averaging 0.70, ranging from 0.61 to 0.76 and the subset scoring the same average of
0.70 and ranging from 0.58 to 0.83. The absence of shadow detail was responsible for the high scores of the subset while sensor noise was clearly responsible for the few low values as the scene became totally immersed in shade.

5. DISCUSSION

Analysis of the correlation coefficients reveals a trend which supports the hypothesis that increased small-scale topography decreases the correlation of imagery of a scene. Figure 17 shows the test scenes ranked in increasing small-scale topography. Correlation coefficients were calculated while topography was made by a qualitative judgment considering the regional as well as local topography rather than just an average of the entire scene.

The overall trend supports the hypothesis that increased small-scale topography decreases the ability to differentiate meaningful edges throughout lighting variations. As the graph shows, the scenes with the least amount of small-scale topography scored the highest correlations. The scene with the least amount of small-scale topography however, the lettering on the trailer, did not receive the highest correlation score. This was most likely because although the scene itself had little real topography and had mostly reflectance differences, features outside of the scene cast shadows across it which fell within the threshold of the edge detection and therefore appeared as topographical features would have. These shadows dropped the correlation from what probably would have been well above 0.60.

The brick wall scene showed the highest correlation of all. Although there was some small-scale texture, it was small enough that any shadows from it did not show up in the image enough to affect the edge patterns. Even the zoomed-in subset had as its lowest a 0.58 which was still higher than the averages of most of the other scenes. The large shadow of the zoomed-out scene affected the correlation very little. It is very likely however that had the images not been processed with the RVS, the scores would have been much lower as noise would have greatly reduced the ability to detect relevant edges.

![Figure 17. Correlation coefficients versus scene small-scale topography. Correlation based on edge images. Topography based on visual assessment.](image)

The scene of the staircase on the side of the liquid nitrogen storage tank was interesting in that there were factors that both increased as well as diminished the potential correlation. In Figure 17, only one representative image is plotted since all of them scored nearly the same regardless of exposure or shadow position. The high relief of the staircase along with the low sun angle created very large shadows which moved great distances across the otherwise evenly-lit and featureless tank surface. This seemed to have had the effect of lowering the correlation as evidenced by the edge patterns of these shadows. Both the exposure and shadow series of this scene showed it to be more heavily influenced by the presence of this source of false correlation of the regular pattern of the stairs and the resulting shadow. It seems that even large shifts in shadow position might result in enough edges lining up to give it an unrealistically high score.
This observation resulted in the additional tests taken from slightly different points of view, thereby intentionally creating different scenes. And although the same scene from other positions scored extremely low, the scores were still above that of what we have been considering to be a “match”, indicating positive scene recognition which was clearly not the case, at least for the purposes of our work. Figure 13. is a set of two edge images of the tank taken from two locations – the location of the first set plus one taken from a position slightly to the left. These edge images are shown in gray. The correlating edges are shown as lightness sums and are shown as white. The correlated edges primarily occur in the areas of high line density, either among the stairs themselves or the shadows of the stairs. With such a high edge density, it's very likely that some edges will coincide, creating an apparent correlation. Slight shifting of the registration is likely to cause major changes in this correlation as more or less edges coincide. This illustrates that regular, dense patterns are likely to reduce the uniqueness of a scene.

The zoomed-out view of the aggregate stone steps of Figure 6 fell about mid-way topographically in this collection of scenes. It's zoomed-in counterpart scored considerably lower due to the larger presence of shadow detail which changed as the sun angle changed. This scene was composed of the front face along with the top edge which lost focus and therefore fine detail in the background, limiting the effect of the shadow detail. Topographically, it was a mixture of some fine detail mixed with an area of smoothness. Although the stones were indeed small, the zoomed-in subset captured much of the detail of the shadows, making this scene much more topographic and thereby receiving a score of about half that of its larger counterpart.

The white wall scene of Figure 7 scored very similarly to that of the aggregate steps. The scene consisted of a rough section taking up about one third of the scene along with a smooth remainder but with deep grooves. These grooves provided shadows which moved as the sun angle changed. The zoomed-in subset, like that of the aggregate steps, consisted of rough texture with fine detail which lowered the score considerably.

The light pole scene of Figure 5 was composed of some very heavy topography and with the sun from the side, created some very heavy shadowing which changed the scene considerably as the sun angle changed. Additionally, there was a fair amount of small-scale topography in the rough surface to affect the correlation considerably. The scores started out relatively high, 0.58, but dropped rapidly as the sun cast large shadows across the smooth surfaces and dropped to a low of 0.19.

The disparity of scores in the tree bark scene of Figure 5 was the greatest of all the test scenes, varying from 0.32 for the zoomed-out scene to 0.14 for the closeup. The RVS processing performed well by bringing out many edges which would have been otherwise lost in the shadows. Although the detail was great in both views, the highly irregular pattern only allowed the occasional correspondence of edges rather than the large correspondence as with the stairs on the nitrogen tank scene. And of course the zoomed-in view contained so much fine detail that although it did score high enough to be considered a match in all but one case, the scores overall were so much lower than any of the other scenes. This followed the premise that increasing small scale scene topography increases the available edges but reduces the coincident edges thereby reducing the ability to recognize a scene as unique from others similar to it.

6. CONCLUSIONS

A correlation analysis was presented which quantifies the degree of pattern constancy and the impact of RVS image enhancement processing on achieving resilient pattern recognition and the dependency of scene context.

Rich topography at the regional and local level of images does result in lower correlation coefficient values due to the prominence of shadow effects in the image data which do translate into more mutable edge data as a significant component of the frame edge pattern. This does lower pattern constancy to a degree, and therefore makes scene pattern recognition more challenging for scenes rich in regional and local topography as compared to scenes that are low in topography content.

The overall result relates to aviation safety external hazards by quantifying that the pattern recognition of terrain hazards is going to be more challenging than runway hazard detection where topographic effects are negligible.
This simply means that reliable scene pattern recognition of terrain hazards will likely be limited to range of higher signal-to-noise ratio imaging conditions than runway hazard detection. Otherwise, edge detection is a viable tool for scene identification for smart vision systems.

7. FUTURE WORK

It might be interesting to compare slightly different scenes to determine the limit of scene recognition over differing points of view. This would be useful for understanding how effective scene recognition might be for automating landing approaches, especially in poor visibility conditions.

For the purposes of this work, it would be useful to conduct a study to determine the very best correlation possible in very controlled conditions to understand the limitations of such a system. Such control would encompass consistent lighting with no changes in illuminant position, absolutely no movement in the camera so as to insure exact image registration, no changes in temperature or vibrations which might affect registration as well, and the use of cooled sensor arrays to reduce sensor noise.

A new edge detection scheme, which is being developed and described in a companion paper by Zia-ur Rahman, “Adaptive methods of two-scale edge detection in post-enhancement visual pattern processing” differs by the way in which it floats the edge thresholds. The new method handles noise and texture removal as a comprehensive issue – assuming that both high noise and dense texture that are denser than printed text are not visually significant and should be removed. This retains weak features such as those which might be in images relevant to aviation safety, i.e., such as a tower barely visible in the distance.

ACKNOWLEDGEMENTS

The authors would like to thank NASA's Aviation Safety Program, Intelligent Integrated Flight Deck, External Hazards for their support of this research program.

REFERENCES


