Chapter 6

Conclusions

This thesis proposes a new approach to represent images. This representation is inspired by emulating an image surface with a rubber cloth stretched by 3-D pipes. The whole procedure is developed with the aid of differential geometry to extract verge points from image surfaces. The proposed verge point representation scheme featured in three aspects: compact data size, efficient feature detection, and flexible image content manipulation.

The extracted verge points with compatible properties are linked into verge curves. These extracted verge points offer an effective way to extract spatial features, like edges, of the original image. In addition, the extracted verge curves can be further compressed using B-spline approximation. Based on the extracted verge curves or the computed B-spline control vertices, a scalable and interactive image codec as well as a flexible image editing architecture are proposed.

In designing a feature-based method for scalable and interactive image compression, the features in an image can be revealed at the decoder side by transmitting the B-spline control points. We construct a hierarchical data structure to encode these control points. The visual performance of the proposed method is comparable to JPEG using the proposed data structure. The control points can be rearranged to support diversified scalabilities, such as SNR scalability, spatial scalability, and shape scalability. The proposed method is suitable for “region of interest” selection and transmission. Experimental results are provided to demonstrate the scalability and interactivity of the proposed image codec.

To detect features on image surfaces, the Hessian operator is used to find edge and corner pairs. For an edge, we detect the edge center, edge contrast, and edge span using high-curvature pairs. The profile of an edge can be approximately reconstructed using these three elements. Meanwhile, the detected edge contrast and edge span can be adjusted for different purposes, such as contrast enhancement and sharpness enhancement. The Hessian tensor can also be used to find corners, which are useful for feature selection and edge linking. Experimental results have shown that the proposed edge and corner operator is comparable to Canny operator in detecting edges and Harris operator in detecting corners. In addition, we demonstrate the feasibility of using verge points to perform region segmentation.

Based on verge point representation, a flexible framework for image editing is also presented in the thesis. An original image is decomposed into the skeleton component and the residual component. By manipulating the skeleton component and residual component, different visual effects in contrast adjustment and sharpness enhancement can be achieved. With the aid of the skeleton curves, we can edit the features in an image adaptively, a difficult task to achieve using conventional methods such as histogram equalization and unsharp masking. By changing the position of the control points used to approximate verge curves, the shapes of objects in an image can be modified. Experimental results have shown the versatility and flexibility of the proposed scheme.

Although the proposed representation scheme is suitable in representing image with sufficient smooth regions or long curves, it may not be efficient in representing texture images. The image surfaces around texture regions vary so rapidly that we need lots of verge points to record texture information. However, we may improve the efficiency in representing textures in two ways. One way is identifying the texture regions first, and then encoding these texture regions using existed texture coding scheme, such as the wavelet-based texture coding scheme used in MPEG4. The other way is downsampling or downsizing the control points in texture regions. After this procedure, the amounts or relative positions of control points are reduced. The coding efficiency can thus be improved.

If we downsize an image, the reduction rate of feature length would be lower than the reduction rate of total pixel number. For a diagonal line in an NxN image, if the image is downsized to N/2xN/2, the reduction factor of length is $\sqrt{2}/2$, while the reduction factor for the total pixel number is 1/4. This implies that the proposed verge point representation is less efficient in representing images in smaller size. Nevertheless, for a diagonal line in an NxN image, if the image is enlarged to 2Nx2N, the incremental factor of length is $\sqrt{2}$, while the incremental factor for the total pixel number is 4. Therefore, if we approach this problem from an opposite point of view, the proposed method may offer good compaction capability in representing large-scale images.

The verge point representation may be extended to represent higher-dimensional signals or videos. To represent 3-D signals, the surface may be considered as (x,y,z,I(x,y,z)). Similar to the 2-D case, we may use 3x3 Hessian tensor to estimated curvatures for a point, and use the largest one in magnitude of the three eigen-values to detect high-curvature points. To represent video data, we may perform verge curve linking in each frame first, and then perform inter-frame curve mapping. For mapped verge curves in the next frame, we record the shape change of these curves. For unmapped curves, we consider them as new nodes, and encoded the verge points or control points in these nodes.