Qualitative Aspects of Chromo-Stereoscopy for Depth Perception*

Thierry Toutin

Abstract

The display of three-dimensional (3D) quantitative data sets is a basic topic of research in cartography, image processing, and applications related to spatial information. A new application for data visualization and analysis, which combines color vision and depth perception, has been developed using the effect known as chromo-stereoscopy based on Einthoven's theory.

It enables the generation of flat color composite images from multisource data in which depth information is coded into colors. When viewed with double prism refraction ChromaDepth™ glasses, a "dramatic" 3D effect is produced. Following a description of the method, the geometric and radiometric processing parameters are qualitatively analyzed to assess their impact on the quality of the chromo-stereoscopic images and depth perception.

Introduction

Through the normal process of vision, we unconsciously evaluate the forms, shapes, distances, and colors of a vast number of objects around us. However, to successfully represent three-dimensional (3D) information on a flat surface such as an image or painting, both the perceptual and conceptual levels of understanding spatial relationships must be combined. Illusions of perceptual space can be generated by use of the linear perspective system, taking advantage of the viewer's conceptual knowledge of the perspective phenomena.

Psychological research has indicated that performance at searching a display is much improved if one knows something beforehand about what is to be looked at (Smith, 1962). Because psychological factors play a major role in perception, the remote sensing expert can "go beyond the information given" in the display of an image. Thus, a viewer with some *a priori* knowledge of the data and of the terrain and with a good understanding of the processing has a more qualitative experience. It has been suggested (Hoffman, 1990) that researchers might devote more time to studying and integrating these qualitative aspects of the remote sensing process.

Chromo-stereoscopy offers a tool to qualitatively perceive depth. Differently colored objects at the same viewing distance can often appear to lie at different depths (Einthoven, 1885). This apparent depth difference can then be enhanced by a refraction process. Therefore, these two known phenomena can be used to perform fusion of multi-source remote sensing data, and to qualitatively perceive depth from the single stereoscopic fused image. The method is straightforward: depth is coded into colors, and then decoded by means of basic optics to produce depth perception. When viewing a color-encoded composite image using chromo-stereoscopy, the stereo effect is based on physiological and psychological factors of color and depth.

The first attempt to produce a chromo-stereoscopic image from remote sensing data (Toutin and Rivard, 1995) resulted in an interesting 3-D product with some artifacts. Published as a highlight article in the October 1995 issue of *PE&RS*, it briefly described the process of depth perception, the chromo-stereoscopic method, and the generation of the cover page image. There was no detailed development of these different aspects, nor any analysis of the relationships between the processing steps and the parameters.

In order to better understand the different components (eyes, glasses, images) of the chromo-stereoscopic method, this paper will expand on the earlier article and address the different factors (physical, physiological, psychological) which influence the generation and the perception of chromo-stereoscopic images. This method examines depth perception qualitatively; it does not extract quantitative depth information which can be directly obtained from the input data. Therefore, only the qualitative aspects of the process are discussed. Because the eye and the brain are part of the "sensor" which acquires and perceives information, some key aspects of their important role in this method will be examined. After reviewing stereo-viewing processes applied in remote sensing and describing in detail the eye as a sensor, the chromo-stereoscopy method and the ChromaDepth™ glasses, developed by Steenblik (1986), are examined. Finally, the technique to control the input parameters of the geometric and radiometric processing is analyzed and tested on different images.

Review of Stereo-Viewing Methods

Different methods have been developed to recreate depth perception. This paper is not intended to fully address all three-dimensional imagery techniques. Only those most commonly used, especially in remote sensing, are described. Okoshi (1976) is a good reference for more details. Depth perception can be "natural," with two images taken from different view points, or "synthetic," to generate a stereo-pair or a perspective view. For the latter, depth perception can be related to any theme such as terrain elevation, magnetic or gravimetric fields, etc.

In natural depth perception, stereoscopic pictures are viewed separately by each eye: the right viewpoint image to the right eye and the left viewpoint image to the left eye. An experienced observer of the stereo-pair may be able to achieve the proper focus and convergence without special

*The ChromaDepth™ glasses, provided in the October 1995 issue of *PE&RS*, can be used to perceive the 3D effect on the chromo-stereoscopic images.

Canada Centre for Remote Sensing, 588 Booth Street, Ottawa, Ontario K1A 0Y7, Canada.

Photogrammetric Engineering & Remote Sensing,
Vol. 63, No. 2, February 1997, pp. 193–203.

0099-1112/97/6302-193$3.00/0

equipment; however, some devices used ordinarily allow each eye to see the appropriate picture of the pair. The device used depends on how the stereo pair was generated. Most commonly, they are conventional stereoscope, anaglyph, or polarized glasses. Recently, 3D shutter glasses are a more sophisticated device used with a computer screen, where shutter glasses are synchronized with a screen alternatively displaying left and right images.

As reported by Okoshi (1976), the first attempt at a stereoscopic drawing was a technique devised by Giovanni Battista della Porta around the year 1600. Today, stereo-paintings have been created by depicting left and right perspectives, as Magritte with his "Man with a Newspaper" (The Tate Gallery, London, England, 1928), or Dali with his "Christ of Gala" (Museum Ludwig, Cologne, Germany, 1978).

In the same way, many proposals for 3D movies have appeared since the first stereoscopic polarized movie in 1939: some relied upon parallax effects with simultaneous projections on the screen; others, like "Cinerama" in the 1950s, relied upon psychological illusion to give a strong, realistic 3D sensation by surrounding the spectators through a wide angle projection.

More recently in 1995, Sharp Laboratories researchers in Europe claimed a major advance in 3D moving image technology (Toutin and Rivard, 1995). To produce the 3D image, "twin-LCD (Liquid Crystal Displays)" are placed at right angles to each other. The images are then combined by a proprietary optical filter that transmits the image from one display, and reflects the image from the other, producing a 3D image. A sensor on top of the device detects the position of the viewer's head on which a tiny silver spot has been placed. The images are then updated on the screen in such a way that they appear to be at a constant angle of $\pm 20°$. Between these limits, the system provides the ability to "look" around the objects without image flipping (Walko, 1995).

But most of these stereoscopic devices have a common feature: the presentation of a separate view to each eye with horizontal binocular parallax. Jones *et al.* (1984) developed a new 3D imaging technique which employs a modified alternating-frame system. With a two-source system, VISIDEP™ utilizes the vertical disparity parallax to generate the depth effect. The two cameras used to generate the images are optically aligned on a vertical plane. Jones *et al.* (1984) discovered the apparent superiority of vertical parallax to horizontal parallax with alternating frame technology. Its applicability is apparently supported (Imsand, 1986). This VISIDEP™ system seems limited to a specific imaging system, but according to McLaurin *et al.* (1986), almost any situation with two sources of information can be adapted to the system. Some applications with remote sensing and geographic data have already been realized with this system (Ursery, 1993).

All the 3D imaging techniques described previously relied upon only two cues: binocular parallax and convergence. The invention of holography (Gabor, 1948) was particularly significant because it offered for the first time a method of spatial imaging with the accommodation cue. Furthermore, it is also the only true three-dimensional photography (Friedhoff and Benzon, 1991).

Holography is a two-step lens-less imaging process: recording an interference pattern on a transparency (the hologram), and reconstruction of the developed hologram. Stated simply, if light A interacts with light B to form an interference pattern that is captured on photographic film, it is possible to project light A onto the developed film, the hologram, to recreate light B. If light B was reflected from an object, then its reconstruction is a holographic image of that object. In an ordinary photograph, what is recorded at a point is the brightness of the image, or the amplitude of the wave reflected by the object at that point. However, in holography, the key point is that the interference pattern contains, by way of the fringe configuration, information corresponding to both the amplitude and the phase of the wave scattered by the object.

Because holography is the only true 3D photography, a hologram can be viewed from different perspectives just as is possible with the objects. For example, if a magnifying glass is recorded on a hologram, the magnifying glass in the hologram magnifies different objects as we change our point of view (Friedhoff and Benzon, 1991).

But, there are practical disadvantages mostly due to the use of coherent light in the recording and reconstruction steps: objects must be still and recorded only in a darkroom on high resolution expensive plate, and the result is monochromatic (black and white). However, since 1964 the use of white-light illumination for multicolor reconstruction has overcome these drawbacks (Leith and Upatnieks, 1964; Pennington and Lin, 1965). Holography applied to remote sensing data has been also realized (Benton *et al.*, 1985). They generated white-light viewable holographic stereograms obtained by digital and optical processing of a Landsat-MSS stereo pair which produced a black-and-white 3D image of the Earth's surface.

The Eye as a Sensor

The eye is the organ through which humans acquire knowledge of their environment by virtue of the light reflected from, or emitted by, the objects within the environment. For humans, the information provided by the eyes undoubtedly plays the dominant role in our interpretation of the environment. But the power to integrate the viewed image to recognize its contour, its color, and its relation with other objects indicates that the process of vision does not merely consist of "seeing," but also of "perceiving and understanding" through the central nervous system. The eye, considered as part of the brain, is fundamentally an organizer. The eye/brain, starting with the activity of the retina, is actively building a world of objects. This suggests that *a priori* knowledge is useful for a better interpretation and understanding of the image: to have a clear idea of what to look for, where to look, and how to look (Hoffman, 1991).

The eye and vision are very complex and sophisticated systems, involving physiological, biochemical, neurological, psychological, etc. processes. Its full understanding is outside the scope of this paper; only some elementary aspects of the eyes and vision are addressed to better understand the process of chromo-stereoscopy.

The eye has two kinds of receptors: cones and rods. The cones are highly concentrated in the fovea and are in direct and one-to-one connection with bipolar nerve cells. The rods are spread in the periphery of the retina in connection with cones. But there is no one-to-one connection with nerve cells: several rods and cones can have the same connection with a bipolar nerve cell. Thus, visual acuity is less precise when using the rods. Cones, which are bigger, are mainly used for detailed acuity in bright light, and also give color acuity. Rods, which are smaller but 18 times more numerous, are used in reduced light, but give only black-and-white acuity.

Color Vision

In everyday life, familiar things are described by particular colors, and these often are identified by references to familiar things: the blue of sky, the green of grass, the red of fire, etc. More scientifically, color theory begins with Isaac Newton who articulated the fundamental problems that still challenge theorists today. In 1704, he described a number of experiments that revealed "puzzling" properties of colors: using a prism, he separated sunlight into the spectrum of colors. Light of different wavelengths is refracted by varying

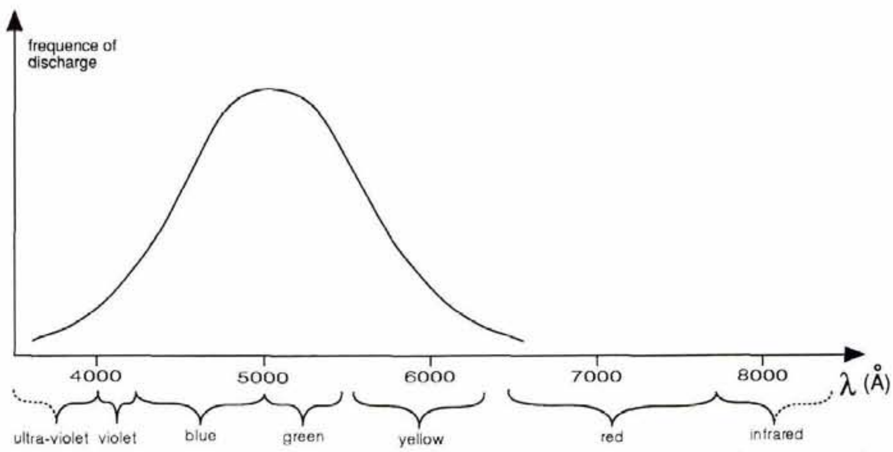

Figure 1. Theoretical wave length response curve of a rod or group of rods measured in frequency discharge in the bipolar cell (from H. Dawson, *Physiology of the Eye*).

degrees when passing from one medium (e.g., air) to another (e.g., a crystal prism). He distinguished seven colors: violet, indigo, blue, green, yellow, orange, and red. These regions represent large ranges of wavelengths (Figure 1). Thus the limits of the visual spectrum are commonly given as 4,000 to 7,600 angstroms (Å).

A color can involve variation in hue (chromaticity), intensity (brightness), or saturation (purity) separately, or it can involve more than one of these components. Within certain regions of the spectrum, subtle distinctions between two hues, as small as 5 to 10 Å, may be appreciated. With extended practice, humans can discriminate 120 or more hues when intensity and saturation are constant, but many more if intensity and saturation vary (Hanes and Rhoades, 1959). In reality, with practical experiments, common experience among scientists showed that differentiation can begin to break down if more than seven to twelve colors along with black, white, and grey tones are used at the same time in a display (McCormick, 1976).

At extremely low intensities (night vision with the rods), the retina shows a variable sensitivity to light according to its wavelength, being most sensitive at about 5000 Å in the green range (Figure 1). In day vision with the cones, there is a characteristic shift to 5,600 Å at the beginning of the yellow range. This has an interesting psychological correlate: with low intensities, red becomes much darker and blue much brighter. As darkness increases, the rod luminosity scale prevails over that of cones.

Depth Perception

Perception, or perceiving, refers to the process whereby sensory stimulation is translated into organized experience. That experience or precept is the joint product of the stimulation and of the process itself. In the "depth" context, the visual system (the process) creates the three-dimensional world (the precept) we experience from the two-dimensional pattern projected onto the retinas (the stimulation). But the fact that we can see depth quite well with one eye closed, or in a photograph or painting, indicates that two eyes are not necessary for a satisfying sense of depth. This dichotomy suggests an intimate relationship between what might be called "object recognition" and perception of three-dimensionality. Unfortunately, at this point we know little about how the brain identifies objects, so a large portion of "depth perception" is simply not understood (Friedhoff and Benzon, 1991). However, in modern psychology, it is accepted that depth perception is based upon as many as ten cues (Braunstein, 1976). The cues have been treated as additional pieces of information which, when added to a flat picture on the back of the eye, make depth perception possible. The brain then combines these cues with the 2D picture to produce judgements about the relationship of objects in space.

These ten cues can be classified into two major groups: four physiological and six psychological obtained from the retinal image (Okoshi, 1976). The physiological cues are accommodation, convergence, binocular disparity, and motion parallax. Accommodation (focusing for distance) and convergence (moving the entire eyeballs) are associated with the muscular activity of the eyes, and they are regarded as minor cues to depth, effective only at a distance of less than nine metres.

The binocular disparity cue is related to the slightly different images of any given object which are received by the eyes embedded at different points in the skull. The degree of disparity between the two retinal images, so-called binocular parallax, depends on the difference between the angles at which an object is fixed by the right eye and by the left eye. For medium viewing distances, this cue is the most important for depth perception. But, at considerable distances from the viewer the angular difference between the two retinal images diminishes, which results in depth perception reduction.

With visual motion parallax, distance cues are obtained from retinal changes that depend on the interposition of objects in space: the rate at which the projection of an object moves across the retina varies with the distance of the object. It is especially significant when the observer or the object is moving rapidly.

The six psychological cues which can give depth sensation from planar images reproduced on the retina being assisted by experience are retinal image size, linear perspective, linear and areal perspective, overlapping, shade and shadows, and texture gradient. They can also be combined together to enhance depth sensation. In remote sensing, perspective is generally the most useful cue and, in radar images, the shade and shadow cue is the most important for depth sensation. More details on these cues can be found in Braunstein (1976) and Okoshi (1976).

Concept of Chromo-Stereoscopy

The concept of chromo-stereoscopy is first based on Einthoven's theory (1885). Chromostereopsis is a phenomenon in

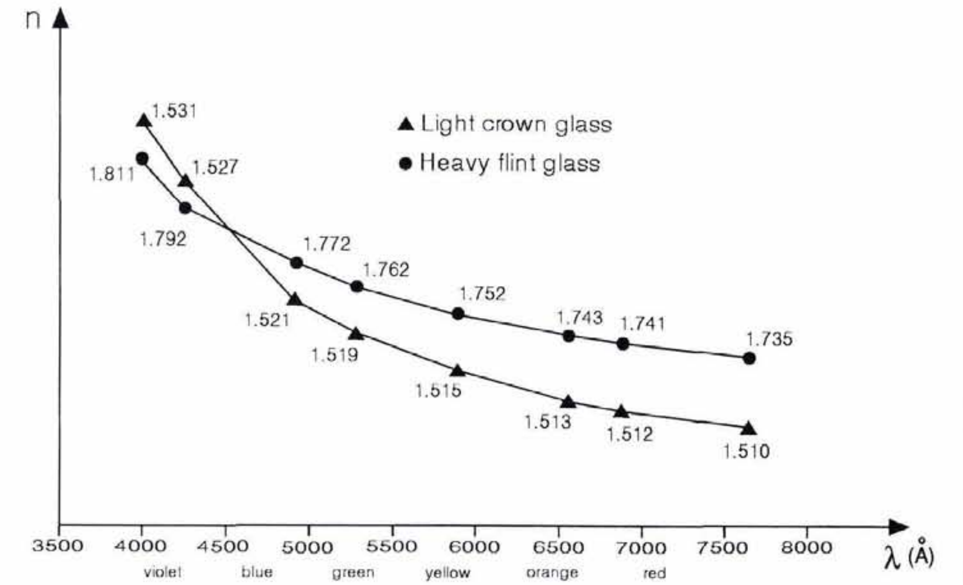

Figure 2. Variation of the refraction index with respect to wave length. The scale for the refractive index is different for the high dispersion heavy flint glass and for the low dispersion light crown glass.

which coplanar colored stimuli are perceived as different in apparent depth. Chromostereopsis is positive if the red object is perceived in front of the the blue one, while the opposite perception gives negative chromostereopsis. Einthoven (1885) credited this phenomenon to transversal chromatic dispersion and the asymetrical relation of the visual and optical axes.

Using the physical properties of refraction, it is then possible to enhance chromostereopsis. A ray of light in passing from air to glass is refracted by an amount that depends on its speed in air and glass, the two speeds depending on the wavelength. For example, upon entering a glass prism, the different refractions of the various wavelengths of the white light spread them apart and each is seen as a band of colored light.

The 3D display technique described in this paper is based on the ChromaDepth™ 3D process and apparatus, developed by Steenblik (1986, 1991), using the two previously described phenomena. Therefore, the basic concept is straightforward: encode depth into an image by means of color, then decode the colors by means of optics, thus producing depth perception.

The first glasses, created with prisms (plastic, glass, Fresnel) and liquid optics, worked extremely well but were not suited for mass production. They later used special optics as a way of making very thin diffractive optics that had the efficiency of refractive optics, and which act like thick glass prisms. Each of the lenses actually incorporates a high precision system of micro-optics, resulting in completely clear glasses.

Each color is then shifted differently, generating differential angular parallax due to the dependence of the refractive index on wavelength. Figure 2 is an example of the refractive index variation for light crown glass and heavy flint glass. But, single prism glasses bend all colors of the light coming from the object, thereby causing the image of the object to appear much closer to the viewer than the actual object distance (Figure 3a). Consequently, it caused two major problems in vision (Steenblik, 1986, 1991):

- Focusing at the actual object distance causes eye strain because the eyes are being focused at a distance which is different from the focal point the brain expects them to have based on the parallax of the eyes.
- If the viewer moves his head while looking at the object, the image of the object moves differently from the expectation of the viewer's brain, causing a visual disorientation or vertigo. This is because the brain believes the object to be nearer than it actually is.

The solution is double prism glasses which push the image distance back to coincide with the actual object distance (Figure 3b). It thereby makes the point in space on which the eyes focus coincide with the point they expect to focus on. The prisms (a low-dispersion and a high-dispersion) are designed so that yellow light entering the first prism passes

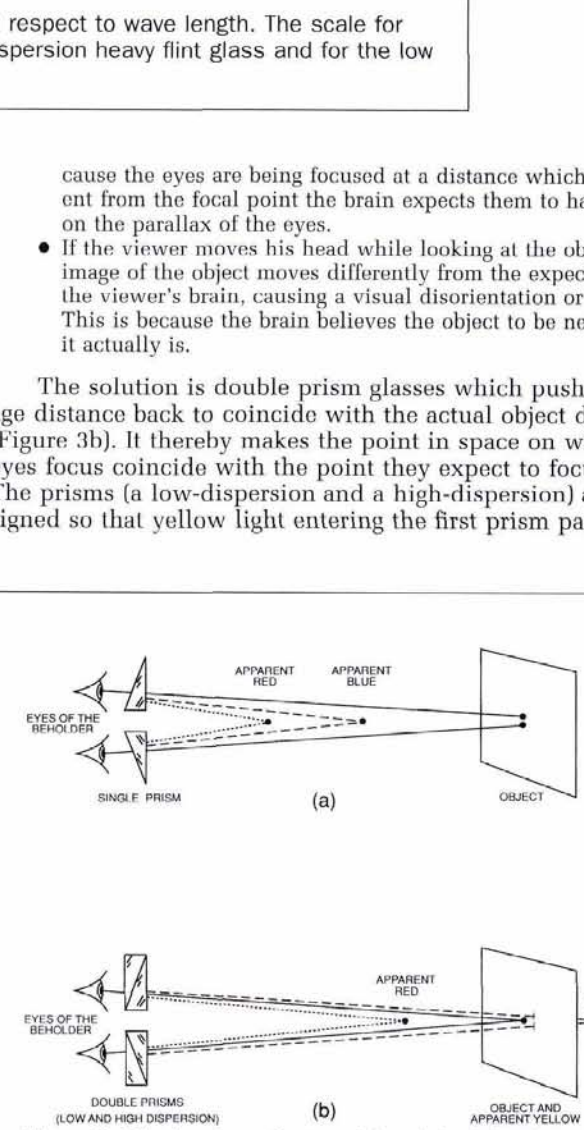

Figure 3. Single prism glasses (a) and double prism glasses (b) configurations for the ChromaDepth™ apparatus (modified from Steenblik, 1991).

through both prisms and emerges parallel to the entering beam. The other colors are shifted with a significant angle deviation due to the high-dispersion prism: red objects appear in the foreground while blue objects are in the background. The yellow image distance and the object distance will nearly perfectly coincide up to a 4- to 5-m viewing distance, without eye strain and vertigo.

Because the eyes are the sensors and the glasses are made with normal prism glass which refracts the light integrally, the physiological characteristics of the eyes in the visible spectrum define the spectral range, the spectral resolving power, and the spatial resolutions of the process. These characteristics have been given in the previous section, "Eye and Vision." Furthermore, the entire process is a tool to qualitatively perceive depth, and the quantitative value for depth can be directly obtained from the data coded for hue to generate the chromo-stereoscopic image. Other characteristics have been given in Toutin and Rivard (1995). They can be summarized as follows:

- a normal flat color image;
- a combination of color vision and depth perception;
- an integration of multi-source data;
- a viewable image from different viewpoints and distances (anisotropic perception);
- a viewable image by fully color-blind viewers;
- a viewable image on print, overhead, computer, laser shows, etc., and;
- a usable image with discrete color on continuous tone imagery.

Because depth information is "synthetically" encoded in color, quantitative or qualitative themes can be used. For quantitative information, color corresponds to value on some quantitative scalar variables (e.g., elevation, gravimetry, etc.). The color coding scheme takes the form of a color-variable to data-variable correspondence. For qualitative information (e.g., polygon data, vector lines, etc.), a palette of colors (differing in hue) is set into a correspondence with the particular characteristics of the symbol (e.g., type of polygon, type of vector, etc.).

Chromo-Stereoscopy Applied to Remote Sensing Data

Why use chromo-stereoscopic images in remote sensing? In the remote sensing process, after the acquisition and communication stages, there is the interpretation stage which involves not only data processing and visualization, but also human interpretation and perception (Hoffman, 1990). Regardless of the level of automatic processing of the data to obtain information, this last stage needs its own specialized tools and systems for the human interpreter to make judgements and decision about the object of study.

Furthermore, color and stereo are the two main characteristics of the eyes: that is,

- for color vision, it has been shown that color can convey information about multi-dimensional data (Ware and Beatty, 1988), and that the use of color can significantly aid users in perceiving information and making judgements (Judd and Eastman, 1971); and
- for depth perception, it has been shown that the interpretation of cartographic information can be facilitated by using 3D or perspective representations, when compared to flat 2D displays (Bemis *et al.*, 1988).

This research suggests that there is increasing interest in stereo and color pictures and how our eyes and brain visualize with machine-handled displays and symbols. Chromo-stereoscopy appears to be an interesting tool to help human interpreters understand remote sensing images and to perceive and extract bio- and geophysical information. But, in order to generate a quality chromo-stereoscopic image from multi-source data which can be "perceived" with the glasses, two separate processing steps (geometric and radiometric) are required to ensure that the chromo-stereoscopic image preserves the geometric integrity and radiometric detail of the input grey-tone images.

As an example, let's use the image published on the cover page of the October 1995 issue of *PE&RS*. Because the study site, the data set, and the processing tools for the geometric and radiometric steps have already been described (Toutin and Rivard, 1995); only a summary and then an analysis of the effects of processing parameters on the quality of the chromo-stereoscopic image and of depth perception are presented in this paper.

Summary

The study site is located near Lake Okanagan in the Rocky Mountains, British Columbia (Canada). The area is characterized by a high relief (1600 m in altitude variation). The data sets are a Landsat TM image, a ground range ERS-1 SAR image, and a 30-m grid DEM generated from the 10-m contour lines of the 1:50,000-scale map. The XYZ cartographic coordinates of 12 ground control points (GCPs) have been acquired by stereo compilation of aerial photographs. The geometric step is an ortho-rectification process with or without DEM depending on the geometric correction model. Radiometric integration is the intensity-hue-saturation (IHS) process with the ortho-rectified Landsat TM3, the DEM, and the ortho-rectified ERS-1 SAR assigned to I, H, and S, respectively.

Geometric Step

Two completely different geometric correction models are used to look at their effects on the resulting chromo-stereoscopic images: a second-order polynomial transformation and a photogrammetric geometric model developed at the Canada Centre for Remote Sensing (CCRS).

This photogrammetric model is based on the collinearity condition and represents the physical realities of the full geometry of viewing. It is integrated because the final equation takes into account the different distortions relative to the platform, the sensor, the Earth, and the cartographic projection. It is unified because it had been adapted to indiscriminately process spaceborne or airborne data, and VIR and SAR data (Toutin, 1995a). Its superiority is mainly due to the fact that it respects the viewing geometry, including relief, and then globally corrects the image. Conversely, the polynomial method minimizes only the residuals at the GCPs (local corrections) without identifying the source of errors. Comparisons between these two geometric correction methods have been addressed quantitatively and qualitatively in Toutin (1995b). This research has shown that the polynomial transformations generate artifacts and erroneous information due to mixed pixels in the co-registration of the input data.

In this study, we qualitatively assessed the effect of the geometric correction method on the quality of stereo-viewing with the ChromaDepth™ glasses. The same parameters are used in the IHS radiometric integration. To better see the details, only sub-images (700 pixels by 1090 lines; 30-m pixel spacing) are presented in Plate 1 (left image from the polynomial method and right image from the photogrammetric method).

As the chromaticity (hue) is controlled by the DEM, the full range from blue to red is presented in both images. The variation of colors is related to different intensities and saturations due to different co-registration of the three data sets.

Without the glasses, the left image (polynomial method) appears more fuzzy, and has a co-registration problem with mixed pixels. Along the west border of Lake Okanagan (marked "a" on Plate 1), the grey tone results from a zero value of the saturation channel because the lake in the ERS-1

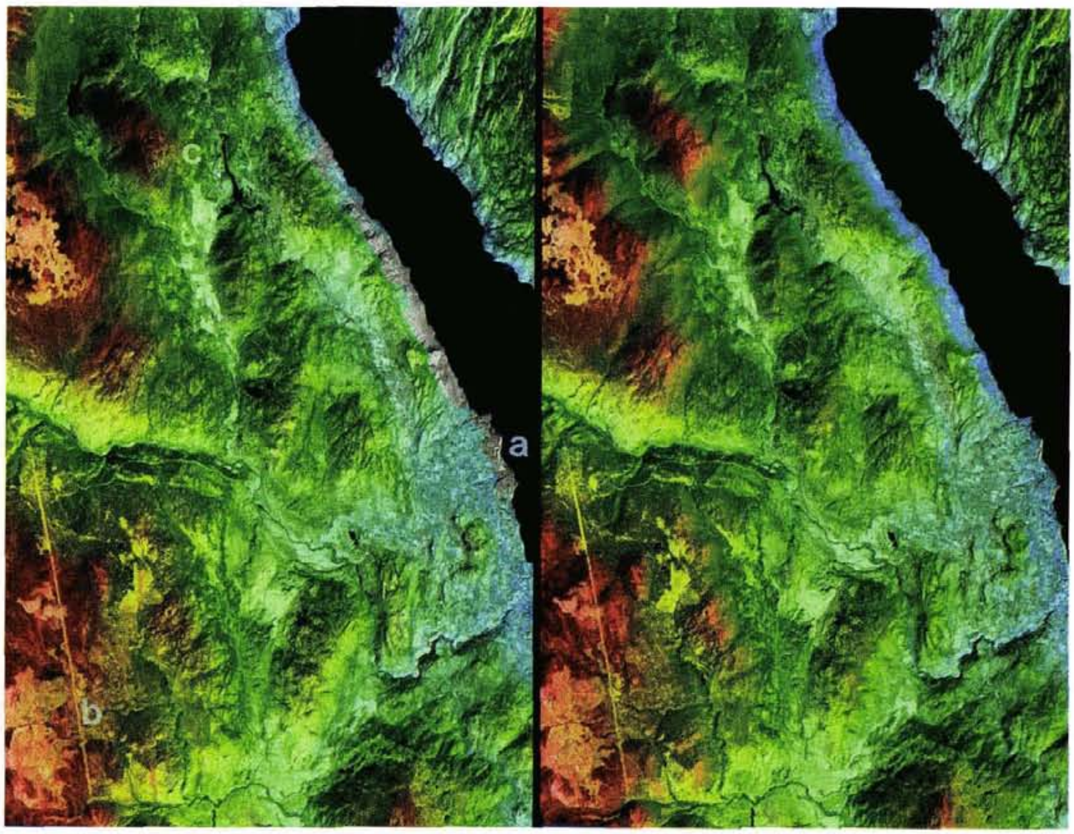

Plate 1. Chromo-stereoscopic images (21 by 33 km; 30-m pixel spacing) generated with a second-order polynomial method (left) and CCRS photogrammetric method (right).

SAR (black in this area) is shifted westward. It is less evident to the north of the lake because there was some backscatter response (maybe wind induced roughness) for the SAR, resulting in different values for the saturation.

Color code variations between the two images are also evident: e.g., the clearcut to the southwest (marked "b") is more yellow in the left image due to a bad co-registration, and the east slope of the mountains to the northwest (marked "c") is more saturated in the right image (CCRS photogrammetric method) due to a better co-registration of the SAR foreslope (brighter backscatter). There is also more correlation between the color gradation and the elevation sensation coming from the Landsat TM3 and ERS-SAR information content.

With the ChromaDepth™ glasses, both images at first glance give depth perception, but the left one (polynomial) is more fuzzy with less detailed depth perception. In some areas (to the southwest), the artifacts generated by the mixed pixels blurred and almost cancelled the depth perception. It seems that there is a contradictory stimulus between binocular disparity "with the color" and the information content from the Landsat TM and ERS-SAR, which the depth perception switched "on and off." The image generated by the CCRS photogrammetric model did not suffer from these problems.

Radiometric Step

In the IHS transformation, one can independently control and vary the processing parameters (mainly the look-up table, LUT) for each component (I, H, S) and see what their effect is on the chromo-stereoscopic image and depth perception. For the first image, published in the October 1995 issue of *PE&RS*, the full dynamic range over 8 bits (256 digital numbers/DN) was used; this resulted in some artifacts in the clearcut areas and at the highest mountain. These artifacts can be controlled, eliminated, or enhanced, depending on the desired effect.

The first artifact is in the clearcut area, which appears to be at the lowest elevation. Because no psychological cue, described previously, plays a role, the depth perception variation between the clearcut and the surrounding forests, which are at the same elevation in the DEM, should be related to a variation in chromaticity. There is then a contradiction.

But the clearcut information comes from the intensity of Landsat TM, which should only change the brightness. Looking at the DN for the three output channels (R, G, B) of the chromo-stereoscopic image, one can see that the DN for the red channel is almost at the maximum value over 8 bits in the forest area. Therefore, a brighter intensity, due to a stronger response of the Landsat TM3 sensor in the clearcut area, changes the color balance, after the IHS transformation, between the R, G, and B output channels because the red channel is truncated at 255 DN. This results in a color with more green and blue then white. Consequently, as white is in the "middle" of the stereoscopic view, the clearcut looks lower. The solution is then to compress the intensity channel LUT to keep the same color balance, after the IHS transformation, between the three channels in these extreme situations. Plate 2 is an example of four different cut-off values to compress the LUT of the intensity channel: from left to right and top to bottom at 200, 220, 240, and 255 DN, respectively.

The variation from "255 LUT" to "200 LUT" results, in general, in a chromo-stereoscopic image which is darker and sharper, with a more realistic depth perception through the

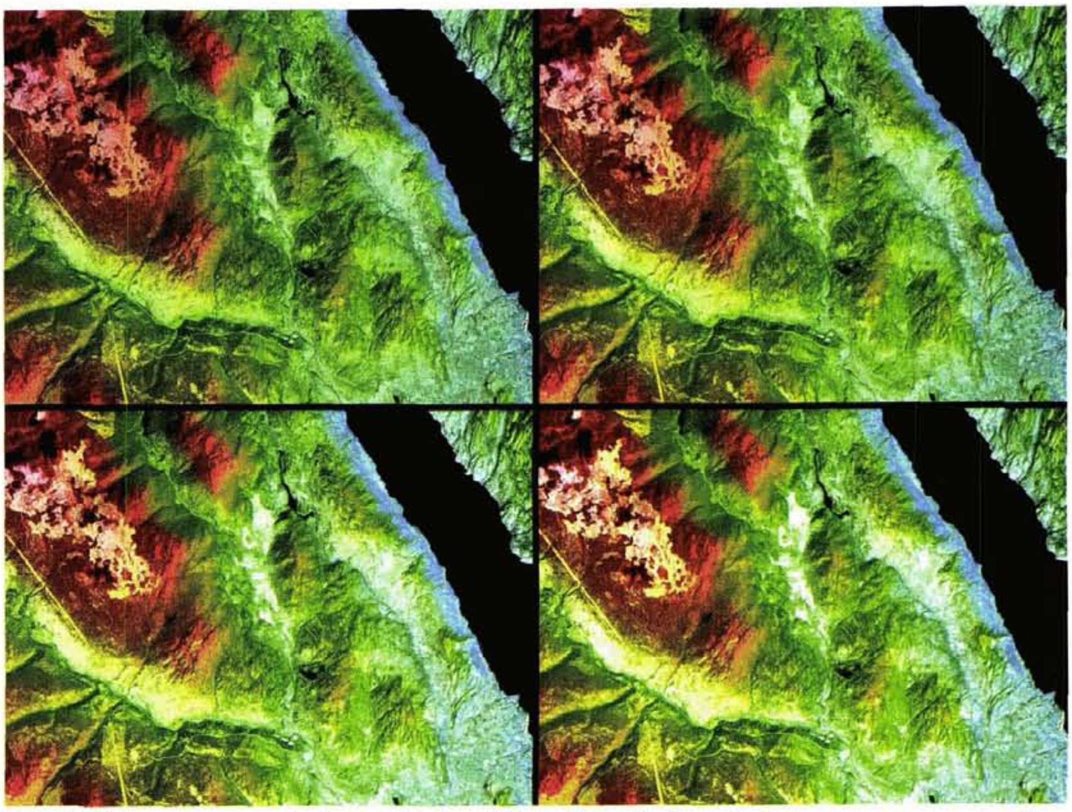

Plate 2. Chromo-stereoscopic images (24 by 18 km; 30-m pixel spacing) comparison by varying the Landsat TM3 LUT parameters of the intensity channel: from left to right and top to bottom at 200, 220, 240, and 255 DN.

glasses due to less color variation between the clearcut and forest area.

The second artifact, which inverted the depth perception at the highest mountain, is due to the violet color. Because in most of the IHS transformation the reference color is blue instead of violet, the shortest wave-length of the visible spectrum, the 0 to 255 DN of the DEM LUT are assigned from blue to blue going through green, yellow, orange, red, and violet (the three primary colors with their basic secondary colors). The solution is then to compress the DEM LUT at approximately three-quarters of the 8-bit range to only use the blue to red range. Violet is not included because psychological on color discrimination suggests using either violet or blue, but not both, and it is hard to discriminate between red and violet (Miller-Jacobs, 1984), as can be noted at the top of the mountain. Plate 3 is an example of four different cut-off values to compress the LUT of the hue channel: from left to right and top to bottom at 200, 220, 230, and 255 DN, respectively. The depth perception at the highest parts of the mountain becomes more and more realistic when looking through the glasses from the "255 LUT" to the "200 LUT" images.

In color coding with the IHS transformation, there is a linear relationship between the elevation value (DN from 0 to 200) and chromaticity (from blue to red) during the coding of depth in the color. But in the color decoding through the glasses, there is not a linear relationship between the color (from blue to red) and the refraction index (Figure 2) proportional to the refraction angle. Consequently, depth perception from the color through the glasses is not linearly correlated with the terrain elevation, even if it is not perceptible qualitatively. The solution is to calibrate the DEM to the specific refraction index variation of the glasses. But unfortunately, one does not have the complete optical properties of the glasses which use double prisms with low and high dispersion materials. Calibration of the ChromaDepth™ glasses' optical properties will be addressed at CCRS in the near future. However, this problem does not seem to be a drawback and thus it does not prevent the use of chromo-stereoscopy as a visualization and depth perception tool in remote sensing.

Finally, the last processing parameter is related to the saturation channel in the IHS transformation. To date, bright, highly saturated hues have mainly been used. But there is no theoretical basis to exclude the use of less saturated hues in a display (Hoffman *et al.*, 1993).

To examine the effect of saturation on depth perception, four constant values (50, 100, 150, 200), instead of the ERS-SAR image, were used for saturation to generate different chromo-stereoscopic images. Plate 4 shows the results with these saturation values from left to right and top to bottom. The saturation increase results in two effects: depth perception increases, but contrasts decrease. When viewing the highly saturated image, it is the cones of the eye, that provide color acuity, that are used. Thus, depth is not more exaggerated but is better perceived in the highly saturated image.

When looking monoscopically, the loss of contrast first gives the impression of losing some details of the Landsat TM3 radiometric content, because the highly saturated image seems "to be burnt." But in stereo with the ChromaDepth™, most of the details are visible, except in the blue area along Lake Okanagan. Furthermore, looking at length at the highly saturated image can cause eye fatigue, but we presently cannot generalize this statement for other images. In general, the images with 100 or 150 DN for the saturation better combine

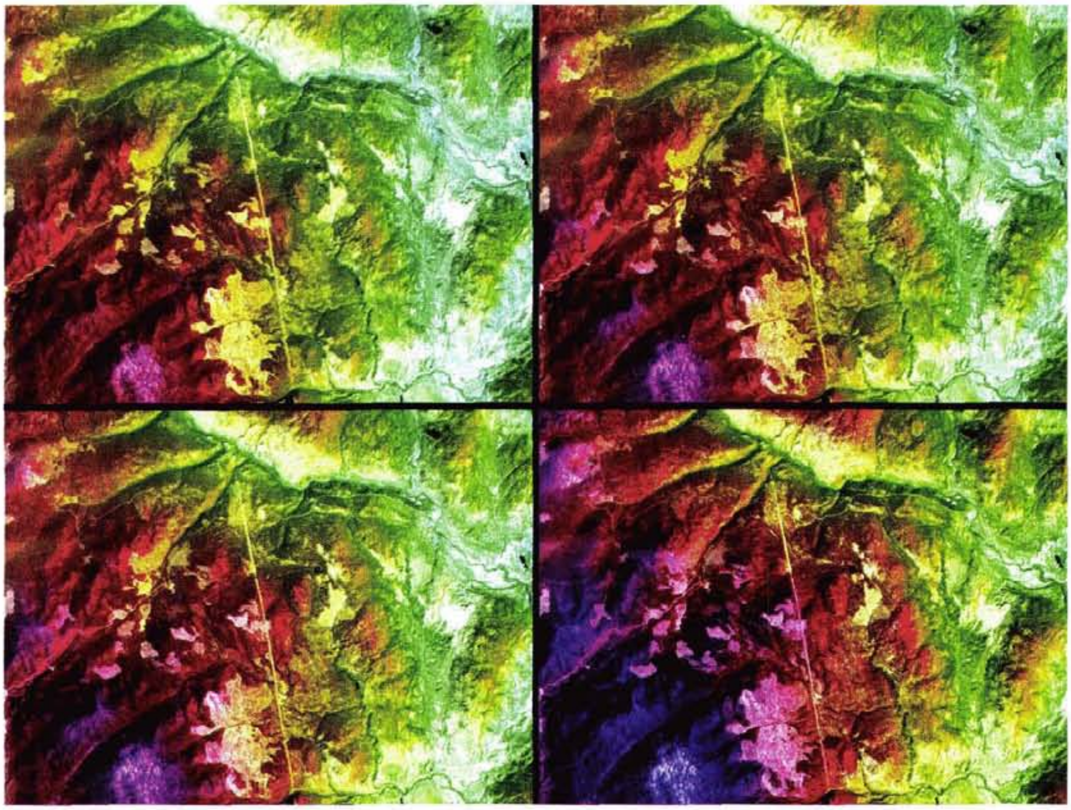

Plate 3. Chromo-stereoscopic images comparison (24 by 18 km; 30-m pixel spacing) by varying the DEM LUT parameters of the hue channel: from left to right and top to bottom, at 200, 220, 230, and 255 DN.

depth perception and the radiometric content of the input data, and are less fatiguing.

Generalization of the Method

The test with this data set and this study site has shown that the choice of the different parameters in the processing steps is dependent of the method: that is,

- the key point in the geometric step is to avoid mis-registration between the different images and the theme; and
- the key points in the radiometric step are to compress the intensity channel LUT, to cut-off the hue channel to only use the blue to red range value, and to have a medium saturation.

Because these parameters are method and not data dependent, they can be applied with a good level of confidence to other data in different configurations. Because it is impossible to show in this paper all the chromo-stereoscopic images tested with these parameters, only the first example is shown (Plate 5), and the others are listed to demonstrate the applicability of the method and its generalization to other data:

- an aerial photo and DEM over Gaspesie, Canada (5 by 5 km; 2-m pixel spacing) at 1:15,000 scale;
- an ERS-1 SAR mosaicking and DEM over Vancouver Island, Canada (500 by 400 km; 100-m pixel spacing) at 1:650,000 scale;*
- an ascending and descending ERS-1 SAR and DEM over the Sudbury Basin, Canada (40 by 40 km; 30-m pixel spacing) at 1:150,000 scale;
- an aerial photos mosaicking and DEM over Ottawa, Canada (40 by 30 km; 5-m pixel size) at 1:50,000 scale;*
- an airborne SAR and a digitized bedrock geology map over Sudbury Basin, Canada (250 by 200 km; 20-m pixel size) at 1:250,000 scale;
- a Landsat TM, or a CCRS airborne SAR and the total magnetic field over Amazonia, Brazil (50 by 50 km, 30-m pixel spacing) at 1:250,000 scale;
- a CCRS airborne SAR and the gravimetric field over Sudbury Basin, Canada (10 by 10 km; 10-m pixel spacing) at 1:40,000 scale;
- a RADARSAT and DEM over Ottawa, Canada (20 by 20 km; 10-m pixel spacing) at 1:80,000 scale;
- a shaded relief of the bathymetry over Atlantic Ocean (100 by 125 km; 40-m pixel spacing) at 1:400,000 scale;
- an aerial photo over Parliament Hill, Ottawa, Canada (1 by 1 km; 0.5-m pixel spacing) at 1:4,000 scale;* and
- a NOAA AVHRR mosaicking over Canada (6800 by 4700 km; 1-km pixel spacing) at 1:8,000,000 and 1:25,000,000 scale.

Because only two input data types are used in these examples, the 150 value was input for the saturation as found in the former test.

The last two examples are extreme and opposite examples in terms of area (6800 by 4700 km versus 1 by 1 km), of pixel spacing (1 km versus 0.5 m), and of scale (1:25,000,000 versus 1:4,000) to show the versatility of this 3D perception tool.

Let's describe the aerial photo over Gaspésie, Canada (Plate 5). The 1:50,000-scale aerial photo was scanned at 1200 dots per inch at the Canada Centre for Topographic Information (Ottawa, Canada). The DEM and the ortho-image were then generated and co-registered with a 2-m pixel spacing on the Heleva/Leica DPW 770 using conventional photo-

*A poster version of this image is available in limited supply from the author.

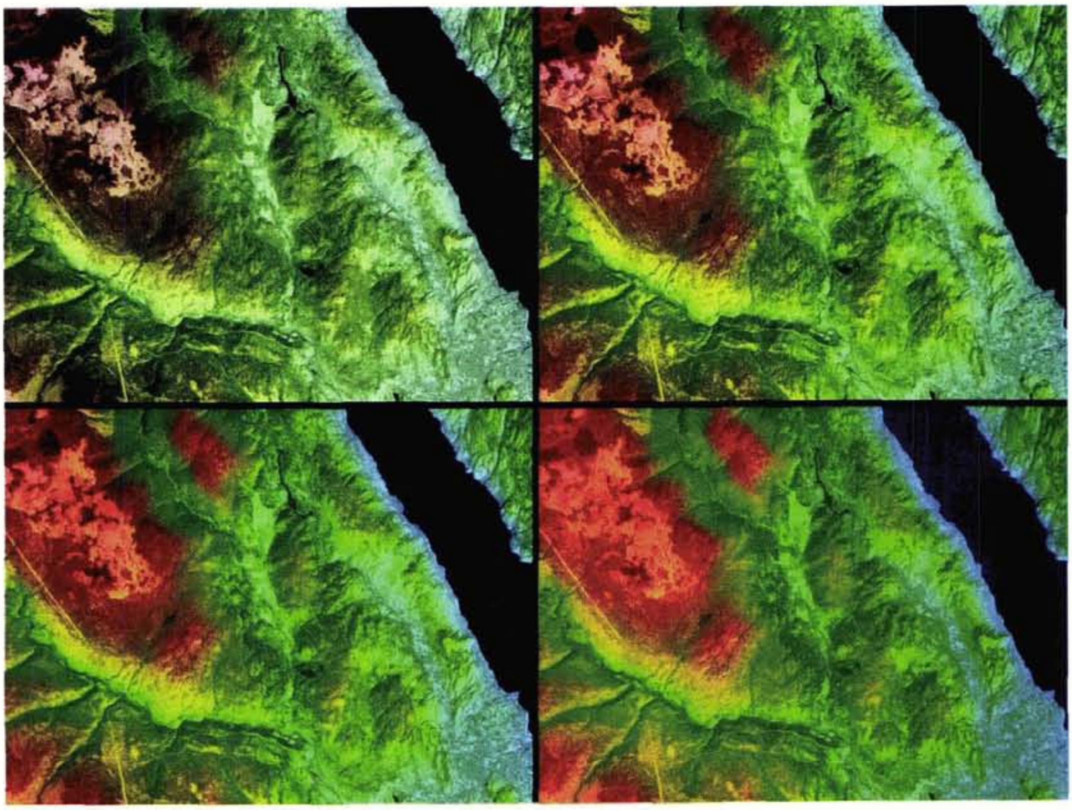

Plate 4. Chromo-stereoscopic images comparison (24 by 18 km; 30-m pixel spacing) by varying the value of the saturation channel: from left to right and top to bottom, at 50, 100, 150, and 200 DN.

grammetric method (Armenakis *et al.*, 1995). The IHS transformation was performed with the same parameters as for the first study site experiment (compression of intensity channel LUT at 200; cut-off value at 180 for the hue and 150 value for saturation).

It can be noted that "white" roads and cities appear above the ground in the blue and green area of the images. The roads and cities (mainly the buildings) are white, the highest DN of the gray range in the aerial photograph. Because it is coded to the intensity channel, it results in very intense blue or green which appear almost white in the resulting chromo-stereoscopic image. As mentioned before, white is in the "middle" of the stereoscopic view, above blue and green in the color depth range. A contrast reduction in the aerial photographs for these features should darken the original color from the elevation coding, and should reduce the artifact.

This two-step process can then be used with most of the 2D data to fuse with a specific theme and to visualize depth with the ChromaDepth™ glasses. If there is a mis-registration problem which generates mixed pixels, a larger pixel spacing can be used to reduce the relative pixel mis-registration. Applicability to data which are already in 3D format is less evident, when combining this 3D perception with chromo-stereoscopy. The different 3D cues (physiological and psychological) should be consistent and generate the same kind of stimuli in the brain to avoid "mis-perception" and interpretation. But if applied carefully with a good understanding of the different depth cues and the impact of the processing parameters as described previously, the same geometric and radiometric processes should give interesting results. More research and tests have to be performed with these 3D data.

Conclusions

This paper has presented a new and complementary tool to display, perceive, and interpret multi-source data. It combines physiological and psychological aspects of the depth perception and color vision processes, which are the main characteristics of human vision. The process which uses the effect known as chromo-stereoscopy is straightforward: encode depth into an image by means of color, then decode the colors by means of refractive double prism optics, the ChromaDepth™ glasses, producing depth perception. As depth perception is "synthetically" generated, any theme such as elevation, gravimetry, GIS polygon data, etc. can be used. The result is a "normal" flat color image which can be viewed without glasses but, when viewed through the ChromaDepth™ glasses, "jumps" into 3D.

Applied to remote sensing data, a chromo-stereoscopic image has to be generated in two steps:

- Precise geometric correction: "precise" is essential because other geometric corrections, such as the polynomial transformation, can blur the chromo-stereoscopic image and its depth perception due to mixed pixels; and
- Radiometric integration using IHS transformation, in which the theme used to generate depth is assigned to the hue channel. The other ortho-rectified images are assigned to the intensity and saturation channels to give the texture and to enhance feature interpretability.

A compression of the LUT of the intensity channel results in a chromo-stereoscopic image which is darker and sharper, and with a more realistic depth perception in areas where large variations in the intensity occur. On the 8-bit scale, 200 appears to be a good value as it has been checked with different data sets. In the same way, a cut-off value

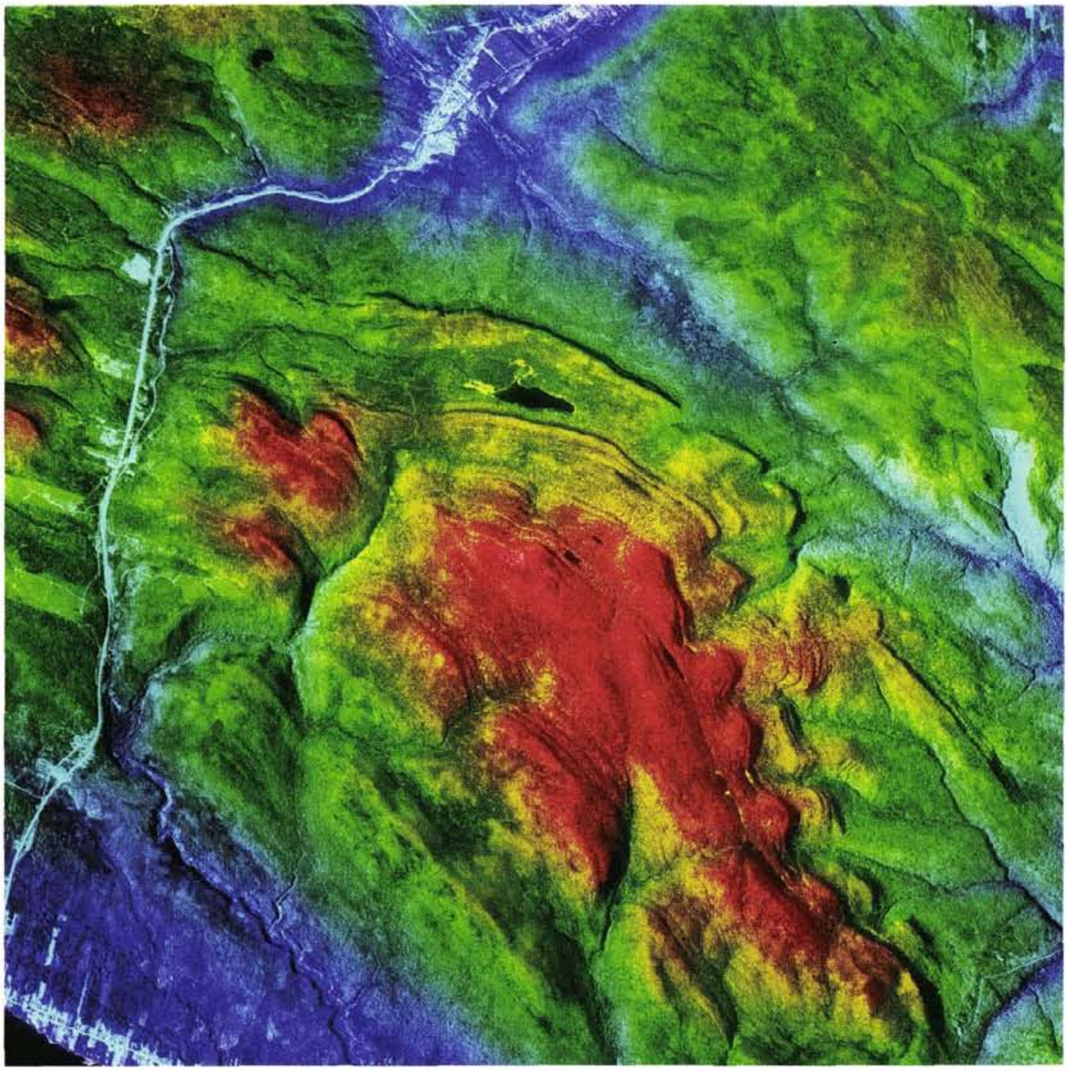

Plate 5. Chromo-stereoscopic image (5 by 5 km; 2-m pixel spacing) of an aerial ortho-photo and a DEM over Gaspésie, Québec, Canada. Saturation value is 150.

(around 200 to 210) of the LUT of the hue channel enables us to use only colors from blue to red. Violet is not used due to a psychological confusion with blue. Finally, if no image is used for the saturation channel, a constant value of 100 to 150 is a good compromise between depth and radiometric content perceptions. It also seems to be less fatiguing than highly saturated images. These parameters are method and not data dependent, which facilitates the generalization of the method to 2D data, and with caution to 3D data.

It has been shown that a good understanding of color vision and of depth perception are essential requirements to enable a better fine tuning of the processing parameters of chromo-stereoscopy to control potential artifacts and to obtain a better depth perception. Depending on the application, the user can voluntarily play on the artifacts with the processing parameters to enhance some specific features as long as there is *a priori* knowledge of the terrain, of the data set, of the method and its applicability. Further examples and applications on the use of ChromaDepth™ will be the subject of an upcoming article. It will include applications in geology, geomorphology, mapping, and environment, as well as use of different data sources (spaceborne, airborne, VIR, SAR, aerial photographs, digital maps, etc.). National atlases and geographic books are also very good candidates for use of this method, allowing to breathe new life into old books. Remote sensing data can be used as a back-drop to these data to add texture and for enhancement of specific themes.

Acknowledgments

The author would like to thank Dr. Benoit Rivard of the University of Alberta (Canada) for the interesting discussions and development on chromo-stereoscopy, and Dr. Frank Ahern and Denny Kalensky of CCRS for the review which improves the paper. The aerial photos over Gaspésie, Canada have been ortho-rectified by Art Dow and Costas Armenakis of CCTI. He also thanks Mr. Martin Coté of Laval University for bringing to our attention the ChromaDepth™ lens. And finally, he acknowledges Chromatek Inc., U.S.A., for providing access to their ChromaDepth™ glasses in support of this research.

References

Armenakis, C., A.M. Regan, and A. Dow, 1995. Softcopy Photogrammetric Applications for National Topographic Mapping, *Geomatica*, 49(4):433–443.

Bemis, S.V., J.L. Leeds, and E.A. Winer, 1988. Operator Performance as a Function of Type of Display: Conventional Versus Perspective, *Human Factors*, 30:162–169.

Benton, S.A., D.A. Duston-Roberge, and R. Simard, 1985. A Chromatic Holographic Stereogram of Landsat MSS Data, *Optical Engineering*, 24(2):338–340.

Braunstein, M.L., 1976. *Depth Perception through Motion*, Academic Press, New York, 206 p.

Einthoven, W., 1885. Stereoskopie durch Farbendifferenz, *Albrecht von Graefes Archiv für Ophtamologie*, 31:211–238.

Friedhoff, R.M., and W. Benzon, 1991. *The Second Computer Revolution: Visualization*, W.H. Freeman and Company, New York, 214 p.

Gabor, D., 1948. A New Microscopic Principle, *Nature*, (161):777–779.

Hanes, R.M., and M.V. Rhoades, 1959. Color Identification as a Result of Extended Practice, *Journal of Optical Society of America*, 49:1060–1064.

Hoffman, R.R., 1990. Remote Perceiving: A Step Toward a Unified Science of Remote Sensing, *Geocarto International*, 5(2):3–13.

———, 1991. Human Factors Psychology in the Support of Forecasting: the Design of Advanced Meteorolgical Workstations, *Weather and Forecasting*, 6:98–110.

Hoffman, R.R., M. Detweiler, J.A. Conway, and K. Lipton, 1993. Some Considerations in Using Color in Meteorological Displays, *Weather and Forecasting*, 8:505–518.

Imsand, D.J., 1986. *Three Dimensional Television Systems*, US Patents No. 4-567-515.

Jones, E.R. Jr., A.P. McLaurin, and L. Cathey, 1984. VISIDEP™: Visual Image Depth Enhancement by Parallax Induction, *Proceedings in Advances in Display Technologies IV* (E. Schlam, editor), SPIE 457, pp. 16–19.

Judd, D.B., and A. Eastman, 1971. Prediction of Target Visibility from the Colors of Target and Surround, *Illumination Engineering*, 66:256–266.

Leith, E.N., and J. Upatnieks, 1964. Wavefront Reconstruction with Diffused Illumination and Three-Dimensional Objects, *Journal of Optical Society of America*, 54(11):1295–1301.

Livingstone, M.S., 1988. Art, Illusion and the Visual System, *Scientific American*, 258(1):78–85.

McCormik, E.J., 1976. *Human Factors Engineering and Design*, John Wiley and Sons, pp. 62–112.

McLaurin, A.P., E.R. Jones, and L. Cathey, 1988. Advanced Alternating-Frame Technology (VISIDEP™) and 3-D Remote Sensing, *IEEE Transactions on Geoscience and Remote Sensing*, 26(4): 437–440.

Miller-Jacobs, H.H., 1984. Colour Displays and the Colour Deficient Operator, *Advances in Display Technology, International Society of Optical Engineering, SPIE*, 457:14–15.

Okoshi, T., 1976. *Three-Dimensional Imaging Techniques*, Academic Press, New York, 403 p.

Pennington, U.S., and L.H. Lin, 1965. Multicolor Wavefront Reconstruction, *Applied Physics Lett.*, 7(3):56.

Smith, S.L., 1962. Color Coding and Visual Search, *Journal of Experimental Psychology*, 64:434–440.

Steenblik, R.A., 1986, 1991. *Stereoscopic Process and Apparatus Using Different Deviations of Different Colors*, U.S. Patents No. 4-597-634 and 5-002-364.

Toutin, Th., 1995a. Multisource Data Fusion with an Integrated and Unified Geometric Modelling, *EARSeL Journal "Advances in Remote Sensing"*, 4(2):118–129.

———, 1995b. Intégration de données multi-sources: comparaison de méthodes géométriques et radiométriques, *International Journal of Remote Sensing*, 16(15):2795–2811.

Toutin, Th., and B. Rivard, 1995. A New Tool for Depth Perception of Multi-Source Data, *Photogrammetric Engineering & Remote Sensing*, 61(10):1209–1211.

Ursery, L., 1993. Virtual Stereo Display Technique for Three-Dimensional Geographic Data, *Photogrammetric Engineering & Remote Sensing*, 50(12):1737–1744.

Walko, J., 1995. Sharp Claims Advance in 3-D Moving Images, *Photonics Spectra*, 29(3):28–29.

Ware, C., and J.C. Beatty, 1988. Using Color Dimensions to Display Data Dimensions, *Human Factors*, 30:127–142.

(Received 19 April 1996; revised and accepted 18 June 1996; revised 15 July 1996)