Compression of Multispectral Images - Combining Spatial and Spectral Compression

Friedhelm König
Agfa-Geveart AG
Munich, Germany

Abstract

One major disadvantage of multispectral imaging is the amount of data that needs to be stored for each multispectral image. In the past, several authors have proposed compression schemes to overcome this disadvantage. Most of these papers dealt with compression either in the spectral or in the spatial realm only, although some authors proposed to divide the data into luminance and color components and to subsample the latter, resulting in a combined method.

Unfortunately, using multispectral imaging the luminance component depends on the illuminant used during the reconstruction of the color. Using a narrowband illuminant every single wavelength could correspond to the luminance information.

This paper proposes a method that overcomes this difficulty by using a more basic approach. The reason, why it is possible to subsample the color channels in conventional color imaging is the dependence of the MTF of the human eye on the wavelength of the observed light. Therefore, it is possible to subsample each spectral band separately (yielding spatial compression). Furthermore, a spectral compression method is proposed, that is compatible to the spatial compression, resulting in a method that combines spatial and spectral compression of the multispectral image.

Finally, the compatibility to conventional imaging will be addressed and it will be shown, that the new method allows achieving limited compatibility.

Introduction

These days, most imaging is metameric. Using this kind of imaging, the spectral information of an original image is reduced to tristimulus information, which is used in turn to produce a copy of the original image. This approach has many drawbacks,[1] and it is extremely difficult to obtain a good copy, especially, if the copy is reproduced using a technology different from the original image or if the viewing conditions are not specified precisely.

Multispectral imaging has been proposed as an alternative approach. Using this approach, as much of the spectral information of the original image as possible is captured and used for the reproduction. Using this approach a precise copy of the original can be achieved comparatively easy. However, the amount of data necessary for an multispectral image is staggering: for each pixel of the image a complete spectrum needs to be recorded. It is easy to see, that this is a major drawback to multispectral imaging.

In the past, a number of papers dealing with the compression of multispectral images have been published. An overview can be found in Ref. [2]. Most of the previous papers dealt with spectral compression only, assuming that spatial compression can be applied to each channel after spectral compression.

Conventional Approaches

One of the most effective methods for compressing conventional images is to split the color information into lightness information and color information, and to subsample the color information. Some authors suggest a similar approach to multispectral information. However, in the case of multispectral images, it is generally not known, which information corresponds to the lightness information, as this is dependent on both the observer and the output illuminant. It is conceivable that the output illuminant consists of only a single wavelength and consequently the lightness information cannot be divided from the color information without making further assumptions.

There are two reasons why this approach is so helpful in conventional imaging. The first is the MTF of the human eye, which is dependent on the wavelength. The second (and more important) reason is the way the human brain deals with color information.

In multispectral imaging, we can only put the first effect to good use by subsampling the original image at each wavelength. The second effect cannot be used with a simple method. Therefore, this paper proposes a compression method that is quite different from the earlier approaches. It is based on an image formation model that is explained in the next chapter.

Image Formation Model

A natural scene consists of a configuration of objects. Each of these objects has a spectral reflectance, which may be different at different points of the object (texture $r(x,y,z,\lambda)$). Usually the color (and spectrum) of a texture changes only slowly between two points of an object. Otherwise it might be assumed, that the object consists of a number of smaller objects (e.g. some floor tiles display such a property). This scene is illuminated by one or more light sources. Each light source emits light with a specific spectral power distribution $I_n(\lambda)$. This light is reflected by the objects and captured within the image plane.

Generally, the light intensity and the power distribution of the light changes for each point of each object I(x,y,z, λ). However, multispectral image applications usually use a light source with known spectral power distribution, as it is desired to describe the scene as if it was illuminated by illuminant E. Accordingly, we can assume that the power distribution of the light changes only in intensity I(x,y,z), as we move from one point of the scene to the next.

Therefore, the light that is captured in the image plan IP(x',y', λ) is dependent on the object texture, the light intensity and geometrical factors, which are given by the three-dimensional structure of the scene:

$$IP(x', y', \lambda) = g(x, y, \lambda) * I(x, y, z) * r(x, y, z, \lambda)$$

For most objects, the viewing angle influences mainly the intensity and not the spectrum of the reflected light (ignoring specular reflections, satin is an exception). Consequently, it is possible to combine the first two factors into a factor I'. The remaining differences can be accounted for by using a slightly modified texture r'(x, y, z, λ). Furthermore, the variables x, y, z can be described as functions of the image plane variables x' and y'. Accordingly the equation can be simplified to:

$$IP(x', y', \lambda) = I'(x', y') * r'(x', y', \lambda)$$

Image Compression

The basic idea behind every compression algorithm is to remove information from an object (image, sound file, text file…) that is not needed because it is either redundant (loss less compression) or irrelevant (lossy compression). In the context of this paper, only loss less compression will be considered (except for the optional subsampling mentioned above). One very commonly used group of methods for loss less compression is based on prediction. Using such a method, the data that has already been decompressed is used to predict the data that has yet to be decompressed. Consequently, only the difference between this prediction and the original data needs to be stored. If the prediction is good, the values of these differences are not equally distributed and an entropy encoding method (e.g. Huffman codes) can be used for an effective encoding of the difference values. As a general rule we can state: the better the prediction, the more efficient the compression scheme.

In the following, this paper will present a prediction algorithm that can be used as basis for this kind of a compression algorithm.

In the technical implementation of a multispectral image acquisition device, the image plane is usually divided into a regular grid of lines and columns. In the following a compression scheme will be assumed that codes the spectra point by point in the order shown in fig. 1. For each point, the complete spectrum is encoded. This is different from many encoding schemes for conventional images. Using such a method, very often, an images is encoded color plane by color plane (that is first all information regarding red, then green than blue). Later on, it will be elucidated, why this is advantageous.

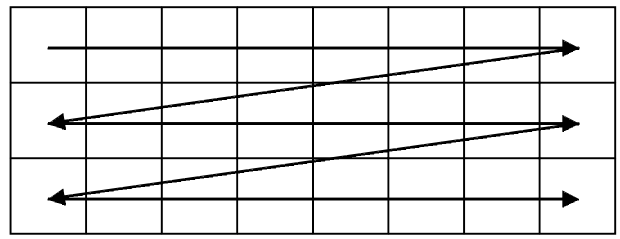

Figure 1. Order used to encode multispectral images

Based on this compression scheme, we notice that all information that is above or left of a point can be used to predict the spectrum of this point. In the following we will use only the points directly to the left and directly above from the point to be encoded (or decoded).

In addition to the spatial order we also need to specify an order for the spectral values. The first idea would be to encode the values in order of increasing wavelength. As we will see later on, another order of the wavelength is more favorable.

Using this order, first the spectral reflectance at a small, a large and a medium wavelength is stored. Next, for each of the intervals the medium wavelength is chosen. This is repeated until all spectral values have been encoded.

For example, if a multispectral image capture device uses the 16 wavelength 400nm, 420nm, 440nm, …, 700nm, the following order could be used: 440 nm 560 nm 660 nm, 400 nm, 500 nm, 600 nm, 700 nm, 420 nm, 480 nm, 540 nm, 580 nm, 620 nm, 680 nm, 460 nm, 520 nm, 640 nm.

Using such a scheme, the prediction algorithm can use spectral information from higher and lower wavelengths to predict a spectral value. Additionally, this allows to achieve limited compatibility with conventional imaging as discussed below.

Methods of Prediction

The prediction algorithms discussed in this paper are based on the image formation model that was discussed above. While examining the spectrum of any given point P_0, we notice that there are two basic situations. Either the points left (P_L) and above (P_A) of this point belong to the same object, or one of these points belongs to a different object.

If all three points belong to the same object, the image formation model discussed above requires, that the spectra of all three points are similar apart from a possible scaling factor that is mostly dependent on scaling. Therefore, we should be able to achieve a good approximation of the given point's spectrum by averaging the spectrum of the two known points and applying a factor based on shading.

As the two known points may have different shading situations, the influence of shading has to be eliminated. This can be done by normalizing and averaging the two spectra $r(P_L,\lambda)$ und $r(P_A,\lambda)$.

$$r_n(\lambda_i) = \frac{1}{2}\,[r(P_L,\lambda_i)/\Sigma_k(r(P_L,\lambda_k)) + r(P_A,\lambda_i)/\Sigma_k(r(P_A,\lambda_k))]$$

Furthermore, a shading factor for the point P_0 needs to be estimated. This can be done by calculating the shading factors for the two known points and averaging.

$$I'(P_0) = \frac{1}{2}\,[\Sigma_k(r(P_L,\lambda_k)) + \Sigma_k(r(P_A,\lambda_k))]$$

Of course, this is usually only a first approximation. However, as soon as some spectral values of the point P_0 are decoded, these values can be used for a much better approximation.

So far, we only considered cases, in which there is a point left and above from the point P_0. Obviously, this is not the case at the left or upper border of the image. In these cases there is only one known point. Here the same method can be applied. But the averaging of the two points has to be omitted.

A different approach has to be found for the upper-leftmost point, where there is no known point at all. In this case, a different prediction method needs to be used. This method will also be used, if not all of the three points belong to the same object. In these cases, only a spectral estimation method can be applied. Luckily, spectra are usually rather smooth. Therefore it is possible to use spectral values that have already been decoded to estimate the remaining values. A very basic approach is to use linear interpolation as shown in figure 2.

Figures 3 and 4 demonstrate the reduction of entropy, that is achieved by the prediction method. Without the use of the prediction method (fig. 3) the code values at 480 nm are almost equally distributed. Using the known spectral values at 400nm and 560 nm the distribution of the code values has significantly changed and the entropy is reduced, resulting in a much better compression. The difference in entropy is even more noticeable, when comparing the distribution with and without prediction at 440 nm, where the spectral values from 400 and 480 nm are available for the estimation method (not shown here).

Obviously other more sophisticated methods of spectral estimation could be used. These methods include, but are not limited to, Wiener inverse, smoothing inverse, Viggiano weighted estimation and many other methods. In general the optimal spectral estimation method to be used, is the one, that minimized the spectral error of the estimation.

Therefore, it is also possible to custom tailor the estimation method to the given image, as proposed in Ref. [3].

As discussed, the proposed compression algorithm uses two completely different estimation algorithms. An important issue is to decide, when to use which estimation method. Basically, there are two different approaches to this question.

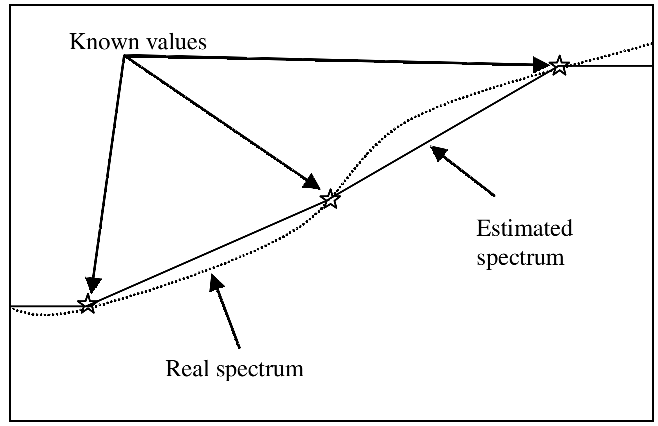

Figure 2. Basic spectral prediction method based on linear interpolation

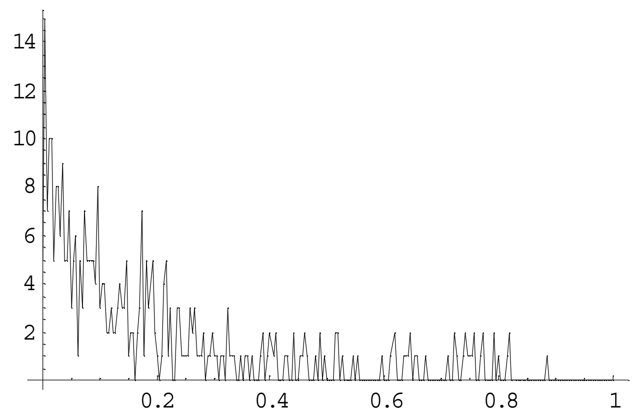

Figure 3. Distribution of code values without estimation method

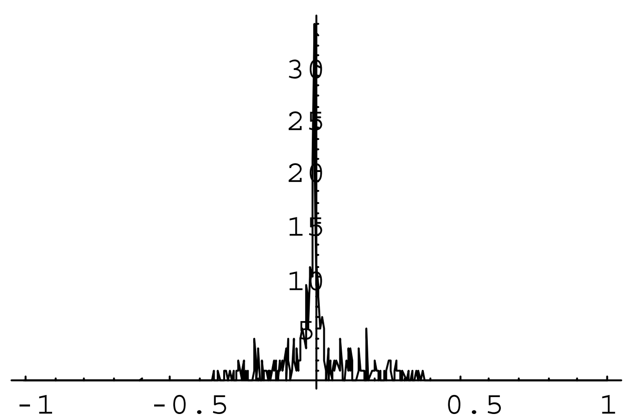

Figure 4. Distribution of code values with estimation method. Notice the reduced entropy.

First it is possible to simply try both estimation methods and to decide which one results in an smaller amount of encoded data for one given pixel. The drawback of this approach is, that the decoder needs to know which method to use. This requires an additional bit for each pixel. However, it is conceivable, that this decision information can be effectively encoded using entropy encoding, because the spatial estimation will be used much more often than the spectral estimation method.

The second approach is to use the encoded data to decide which estimation method to use. In other words, to use the first few (two or three) spectral values to estimate, which estimation method will result in the better compression.

For example, if the first 3 spectral values are reduced by roughly 10% compared with the pixel left of the processed pixel and by 5% compared with the pixel above, it is likely, that the spectrum is basically the same and spatial estimation should be used for the remaining spectral values. If, however, two spectral values roughly stay the same as the spectral values of the pixel left of the processed pixel, and the third spectral value is reduced by 50%, it is likely, that the new spectrum belongs to a different object. Therefore, spectral estimation should be used.

The optimum method to estimate which estimation method to use, is as such dependent on the estimation methods used for spectral and spatial compression and the image to be encoded. Further research needs to be conducted.

Compatibility

One very important aspect of any multispectral imaging system is its compatibility with conventional imaging. Only if multispectral imaging fits into such a system the wide use of multispectral techniques will become viable. Otherwise, the use of multispectral imaging will be limited to the small number of professional applications, in which color precision is of utmost importance, although exact color information can also be helpful for other applications, when a pleasing color reproduction is more important, than a precise color reproduction.

Compatible data formats have been discussed elsewhere.[4] Thiese methods encode an image using a conventional image format (e.g. TIFF) and add additional information in the form of tags or headers. The previous approaches use exact color information with regard to a specified illuminant. This information is calculated from the multispectral information. That is why these approaches do not fit in well with the encoding of spectral values as discussed in this paper. On the other hand, taking a look at conventional imaging, color information usually is not really that precise, due to sensor metamerism and undefined or unknown illuminant situations.

Therefore, does compatibility with conventional imaging really require optimum color values for one given illuminant, which usually is not used to view the image in practice, or is a more relaxed approach viable?

Especially, such a relaxed approach is viable for most consumer application. In this case, it is possible to select three spectral values and to calculate approximate color values from the spectral values by simply applying a matrix transform.

An experiment has been conducted to estimate the quality achievable by this approach. To this end, a set of spectra measured by Vrhel and co-workers[5] has been used. This set contains 354 spectra of natural objects and paints, each sampled at 61 wavelength between 400 nm and 700 nm. The set has been found to offer a very wide range of spectral variance and, therefore, is a good set to use, if the influence of spectral variability is to be estimated.

Tristimulus values XYZ with respect to illuminant D65 were calculated for the 354 spectra contained in this set. Next, three wavelengths were chosen and for each spectrum the sample values at these wavelengths were considered.

Consequently, for each spectrum, there are two three-dimensional vectors, one containing the tristimulus values and a second containing the samples at the chosen wavelengths. A pseudo inverse was used to derive a matrix, that allows the approximation of the XYZ values from the sample values. Due to the nature of the pseudo inverse, this matrix is optimal regarding the quadratic deviation between the original XYZ values and those derived by using the matrix.

Lab values were calculated both from the original XYZ values and from the XYZ values calculated by use of the matrix., and finally ΔE_{94} color errors were derived.

This process was repeated for each of the 43680 meaningful combinations of wavelengths (The wavelengths need to be different from each other and the order of selecting the wavelengths does not matter).

Finally, from these combinations, the combination that leads to the smallest mean ΔE_{94} – error was chosen. The average error in this case was 3.58 ΔE_{94}, while the maximum error was 18.53 ΔE_{94}.

The magnitude of this error is comparable with the error achieved by conventional three-channel imaging devices. Furthermore, this result is not optimal for several reasons. First, the use of XYZ errors for minimization does not correspond too well to the resulting ΔE_{94}–errors. It was chosen for simplicity, as the calculation had to be repeated quite a few times.

Second, a given image usually displays a considerably smaller spectral variability than the Vrhel set. Therefore, it is possible to custom tailor the matrix transform for each image, reducing the color error resulting from this approximation.

Using this approach, it is possible, to encode the spectral image in a compatible way by storing the XYZ (or L*a*b*, or RGB,…) values calculated by the matrix transform in a conventional image format and providing both the transform matrix and the remaining spectral information separately, possible within the same file.

Conclusions

In this paper a new approach to the compression of multispectral color images has been proposed. This method is based on the compression of spectral channels instead of reducing the number of channels to be encoded. It is based on a prediction method, in which for each pixel of the image one of two possible prediction methods is chosen.

This choice can either be automatic or dependent on an additional image plane.

Furthermore, compatibility with conventional image three-channel imaging has been discussed and it was shown, that it is possible to achieve similar color accuracy with this approach. In an experiment an average color error of 3.58 ΔE_{94} was achived.

References

1. Friedhelm König and Patrick G. Herzog, On the limitations of metameric imaging, Proc. of PICS ’99, pp. 163-168, Savannah, GA, USA, 1999
2. Friedhelm König and Werner Praefcke, Multispectral image encoding, Proc. of IEEE’s ICIP’99 International Conference on Image Processing, volume 3. pp. 45-49, Kobe, Japan, 1999
3. Mitchell R. Rosen, Mark D. Fairchild, and Noboru Ohta, An Introduction to Data-Efficient Spectral Imaging, Proc. of CGIV 2002, Poitiers, France, pp. 497-502, 2002
4. Friedhelm König and Werner Praefcke, A multispectral scanner, In Lindsay MacDonnald and Ronier Luo, editors, Color Imaging in Multimedia, chapter 7. pages 63-73. John Wiley & Sons, Ltd., 1999
5. M. J. Vrhel, R. Gershon, and L. S. Iwan, Measurement and analysis of object reflectance spectra, Color Res. and Appl., 19(4): 4-9, Feb. 1994

Biography

Between 1997 and 2001 Dr. Friedhelm König was a scientist at Aachen University of technology, Germany and published numerous papers on multispectral imaging and related subjects. During this time he spend five months at the Natural Vision Research center in Tokyo, Japan, where he was a visiting scientist. In 2002 he published his Phd thesis, which was awarded both the Borchers-Medalie and the Friedrich-Wilhelm award.