Colouring the Near-Infrared

Clément Fredembach and Sabine Süsstrunk
School of Computer and Communication Sciences, Ecole Polytechnique Federale de Lausanne (EPFL), Lausanne, Switzerland
{`clement.fredembach,sabine.susstrunk`}`@epfl.ch`

Abstract

Current digital camera sensors are inherently sensitive to the near-infrared part of the spectrum. To prevent the near-IR contamination of images, an IR blocking filter (hot mirror) is placed in front of the sensor. In this work, we start by replacing the camera's hot mirror by a piece of clear glass, thus making the camera sensitive to both visible and near-IR light. Using a pair of lens-mounted filters, we explore the differences in operating the camera to take visible and near-IR images of a given scene.

Our aim is to enhance the visible images using near-IR information. To do so, we first discuss the physical causes of differences between visible and near-IR natural images, and remark that these causes are not correlated with a particular colour, but with atmospheric conditions and surface characteristics. We then investigate image enhancement by considering the near-IR channel as either colour, luminance, or frequency counterpart to the visible image and conclude that using information from two different colour encodings, depending on the image content, produces vivid, contrasted images that are pleasing to the observers.

Introduction

In contrast to film photography, digital sensors are very sensitive not only to the visible spectrum of 400-700 nm, but also to the near-infrared (IR) part of the electromagnetic spectrum: 700-1100 nm. In fact, silicon is sensitive to near-IR that an infrared filter, "hot mirror", is placed in front of most digital camera sensors to prevent near-IR contamination of the colour signals.

Effectively, this means that by removing the hot mirror, most consumer cameras automatically become near-IR capable. Indeed, this is being routinely done by amateur photographers for specific applications, such as astrophotography or landscape photography. Additionally, while near-IR film existed, its response was very slow, whereas digital cameras have similar exposure time in the visible or the near IR, thus allowing a greater diversity of images.

Our goal here is not to restrict ourselves to near-IR images, but rather to use information from both the visible and near-IR parts of the spectrum to enhance digital images. To do so requires replacing the hot mirror of a camera by an equivalent (in size) piece of clear glass. Using an IR-block or IR-pass filter in front of the lens allows us to capture either a near-IR or visible image of any given scene.

Despite the continued presence of the colour filter array in front of the sensor, near-IR images are effectively grayscale images if white balanced appropriately. It implies that one can consider the combination of visible and near-IR information as a four-channel image that covers a spectrum range of 400-1100 nm. To represent that information in a meaningful manner, one must have a specific application in mind. Ours is to enhance visible images using near-IR inherent properties, yielding images with high contrast, "dramatic" lighting, and more vivid colours.

To do so, we do not consider the near-IR channel as a "fourth colour", but rather as containing spatial information and lightness. We investigate different manners to combine that information with conventional RGB images and perform a psychophysical evaluation to find out which representation observers prefer.

The rest of this paper is organised as follows: in Section 2 we review the different approaches to treat near-IR data that have been proposed to date. In Section 3, we make the case for using near-IR as spatial and lightness information, and review the physical phenomena that motivate our approach. Section 4 deals with image (RGB and near-IR) acquisition parameters that allow us to obtain suitable images for further processing. The experiments and results are presented and commented in Section 5, while Section 6 concludes the paper.

Near-IR approaches

Near infrared imaging has been done for a long time, and many different applications thus exist. Its place in image processing is, however, unique given that a grayscale version of a near-IR image is easily interpreted by a human observer. Contrarily to, e.g., X-ray images or thermal infrared, the scene content is similar to what is perceived in the visible spectrum. Near-infrared images nonetheless lie beyond the visible spectrum, and as such, their colour interpretation is difficult to ascertain. This duality is one of the primary reasons why near-IR applications are generally very task specific.

Near-IR is, however, used in many imaging tasks. A vast amount of research has been devoted to security applications. Indeed, one can easily find manufactured light sources that emit solely in the near-IR, which implies that they are not perceived by the human eye. One can thus capture images or videos in the dark with acceptable quality[1]. Similarly, by fitting such a light source on a camera, e.g., by adding a filter to a flash light, one can obtain photographs that are almost invariant to ambient light sources, an approach which has notably been used in face recognition [1].

Another field that uses near-IR signals is remote sensing, where they are often employed in conjunction with middle- or far- infrared to better distinguish several object classes such as water, constructed areas, and particularly vegetation, where it permits to distinguish between coniferous and deciduous forests [2], [3]. Regarding vegetation, near-IR has also been employed

---

[1] most security companies offer these products on their website.

for in-depth analysis of tree types and health [4].

In a more general setting, an object's near-IR spectral response can be used to discriminate between samples that are visually indistinguishable. Among these are several analysis of counterfeited products such as cigars or drugs [5], disambiguating black metamers [6], as well as discriminating painted surfaces [7]. In all these studies, the visible spectra of considered objects or surfaces were almost identical. While spectra is mostly used in these applications, it was proposed in [8] that a simple digital camera with a 5-filter wheel could be used to approximate near-infrared spectra without significant losses.

The aim we pursue here is, however, different. Near-IR photography has been popular, both in film and digital form, for its sharp, full of contrast, images. The preferred manner of rendering these images is usually in black and white (in most cases the "native" format). Colouring has to be done by hand, whether the end goal is false- or natural-looking colours; and, apart from replacing one channel of the RGB image by the near-IR information, there is little in the scope of an automatic process for obtaining good quality colour-IR images (the main source of information in fact comes from photographers' websites or forums, where people exchange software-based manipulation techniques and the choice of palettes to colour near-IR images).

Near-IR as Spatial Information

In this section, we review some physical phenomena that result in fundamental differences between RGB and near-IR images. Moreover, the majority of these phenomena are not, strictly speaking, colour-related, and so we argue that near-IR should not be perceived as a fourth colour but rather as additional spatial information or lightness.

Rayleigh Scattering

Small particles scatter incident light, which can alter the light intensity at specific wavelengths. When the particles' size is very small ($< \frac{\lambda}{10}$), light behaves according to Rayleigh scattering, which states that:

$$\frac{I_S}{I_0} = \frac{\text{constant}}{\lambda^4} \tag{1}$$

i.e., the intensity of the scattered light I_S is related to that of the incident light I_0 by the inverse of the fourth power of the wavelength λ. As a result, sky appears blue because it is the most scattered colour due to its relatively short wavelength. In comparison, near-IR at 1000 nm is 40 times less scattered than blue at 400nm. We thus expect near-IR intensity to be dramatically lower than its visible counterpart in sky and its reflected components, such as water.

Reducing the influence of Rayleigh scattering does not only make the sky darker. Indeed, one can also enhance the amount of information present in an image when this information has been scattered by atmospheric haze or pollution. Since the decrease of incoming light's intensity is proportional to the depth of the haze and the fourth power of λ, a near-infrared image is by nature less sensitive to its presence. This phenomenon is at times striking, as shown in Fig. 1

Figure 1. On the left, a conventional RGB image. On the right, a near-IR capture of the same scene. The haze has disappeared, revealing a sharper, cleaner, picture.

Mie Scattering

When the size of particles in front of the light path increases, the physical interactions between light and particles change from Rayleigh-type to Mie-type scattering. When the particles are large ($> \lambda$), spherical, and diverse, there is no more dependency between scattering intensity and wavelength, i.e., all wavelengths are equally scattered. This is the physical explanation of clouds' white colour, in opposition to the blue of the sky. From a near-IR perspective, these different relationships imply that while the sky will have very little influence on the near-IR image, clouds should appear unchanged from the RGB image, thus increasing their contrast. This effect is most noticeable in bright days, just above the horizon. An illustration of this effect can be seen in Fig. 2

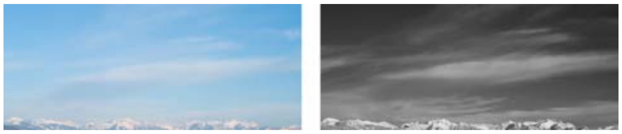

Figure 2. On the left, a conventional RGB image of clouds in the sky. On the right, a near-IR capture of the same scene where the contrast between the different sky elements is greatly enhanced.

Molecular structure: The example of vegetation

The interaction of light and particles in the atmosphere is not the sole reason as to why the spatial content of the near-IR image differs from the visible spectrum. In the case of, e.g., vegetation, differences have to do with molecular structure itself [9]. Chlorophyll gives vegetation its green colour, but it is not all. Its molecular structure also causes vegetation to "glow" when viewed in the near-IR. The intensity of near-IR light reflected by plants varies depending on the season and the type of plants, a much studied phenomenon in biology and remote sensing. In our case, however, we are mostly concerned about the difference in intensity and contrast between the channels. Such an example of vegetation representation is shown in Fig. 3.

Seeing unseen differences.

A common use of near-infrared imaging is to distinguish between surfaces that appear identical to the human eye. Indeed, near-IR is "transparent" to a number of colorants or paints and, as such, it can "see" through the colour layer to reveal the surface underneath. The reverse is also true, there are surfaces that can be distinguished in the visible spectrum that appear the same in the near-IR. An illustration of everyday objects following this

Figure 3. On the left, a conventional RGB image of vegetation. On the right, a near-IR capture of the same scene, note the large difference in contrast between the two images.

behaviour is provided in Fig. 4

Figure 4. A scene with colourful objects. Note that the red and white scarf pattern (middle of the image, green arrow) disappear in the near-IR. On the other hand, the two black objects (left side of the image, yellow arrows) have a very different near-IR response.

Image Acquisition.

We describe here the camera and pre-processing steps taken to ensure that the images, both RGB and near-IR, can be processed in a meaningful manner.

Camera and Filters

The camera we used in these experiments is a Canon 300D, a single lens reflex camera that was modified to make it near-IR ready. We have replaced the hot mirror in front of the sensor by a piece of clear glass with the same size and thickness, thus preserving the focusing capabilities of the camera. This modified camera captures visible and near-infrared light at the same time; we have to use lens-mounted filters to capture the two different images.

The two filters used to selectively take an RGB or near-IR image are a B&W IR-cut filter and a Hoya R72 visible-cut filter that starts transmitting light at 700nm. In effect, images that are taken in the visible light cover a spectrum ranging from 400-700nm, while the near-IR ones range from 700nm until the wavelength is greater than the sensor intrinsic sensitivity (usually around 1100nm, depending on the sensor). The filters are mounted onto a 50mm fixed focal length Canon lens and are manually swapped between each shot.

Illuminants and “White Balance”

Despite removing the hot mirror from the camera, the Bayer colour filter array is still present in front of the sensor, thus allowing the camera to continue capturing normal RGB images; the near-IR data, however, should be monochromatic. For that to happen, however, the CFA filters must be equally transparent to near-IR wavelengths. To assess that, we first have to manually white balance the images. Indeed, none of the preset modes that can be found on the camera are adapted to representing light in the near-IR, besides, it has been reported that between 700-780nm the blue and green filters have a tail absorption and as such a red colour cast is likely. The rationale is that this absorption tail induces a different near-IR quantum efficiency for the three filters, due to their different transmittance characteristics. After white balancing the images, using either a white ceramic tile or a macbeth colour checker, this colour cast disappears and one is left with a proper grayscale image, demonstrating that the presence of the CFA does not hinder the capture of near-IR images, as illustrated in Fig. 5.

Figure 5. Near-IR image with automatic (left) and custom (right) white balance. The properly white balanced image is, straight out of the camera, effectively a grayscale image despite the presence of the CFA in front of the sensor.

The choice of illuminants is also important in obtaining comparable near-IR and visible data. The most often used light sources were sunlight, studio incandescent lights and the camera built-in flash. In general, office lighting (and most gas discharge lamps) alone should be avoided given that their emissions in the near-IR is very low and the resulting images therefore are noisy. In such cases, we often made use of the camera flash, which provides ample visible and near-IR light. Following the same reasoning, near-IR only lights, such as some type of LEDs, were also not employed.

Focus and Exposure Times.

To adequately compare near-IR with visible images, they have to have a similar overall sharpness and brightness. Since adding an IR-pass filter effectively prevents the camera user to see anything and removes the camera auto-focus’ capability, one has to estimate the correct focus point. Considering that light refraction is wavelength dependent and that most lenses are not achromatic in the near-IR range, we have run tests in the lab that tell us how much the focus should be changed from the visible image one so that the near-IR image remains sharp.

Rigourously calculating the proper exposure time of the near-IR image is more complicated. Indeed, a visible high dynamic range image is not necessarily so in the near-IR and vice versa. For general scenes, we have computed compensation factors depending on the light source and distance. Should the image, upon capture, be over- or under-exposed, the settings can be manually adjusted. In most cases, further adjustments, however, were not necessary.

Image Registration

As we will be comparing two images of the same scene, we have too look at image registration, since the two images will not, in general, be perfectly aligned. To automatically realign the images, we use a combination of keypoints matching and homographies. In a first step, we use the SIFT detector to compute keypoints and their correspondence on both the visible and near-IR images [10]; since the images depict the same scene, the matching is generally robust. In a second step, we use Peter Kovesi's Matlab functions[2] to calculate an homography that best explains the movement between the two images.

This method works well, but is unable to assess 3D differences such as a foreground motion that is different from the background motion. In most tested cases, however, it compensated the rotations and translations in a satisfactory manner.

Experiments

The goal of our experiments is to find out whether near-IR can be thought of as lightness or spatial information and can be used to enhance visual images. 20 pairs of images were analysed, a few of them shown here (the other ones are available online[3]), depicting various scenes and objects. All the experiments have in common that the input data contains four channels (R, G, B, and NIR), while the desired output has only three; our goal is to find out which method performs best from a visual standpoint -with the purpose of creating more "pleasing" images. We present here results for each method.

Colour swap

Perhaps the simplest method to display visual and near-IR information at the same time is to simply consider near-IR as an extra channel. One can therefore swap this fourth channel for one of the original ones, creating an illusion of a regular image. Unfortunately, this method is too simple to provide meaningful results given that near-IR is not really correlated with a single colour channel; the resulting images can therefore be awkward, as shown in Fig. 6, which indicates that one should indeed consider near-IR data as spatial information or lightness rather than colour.

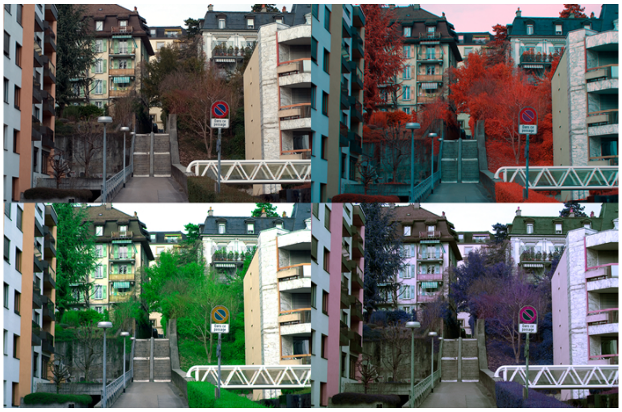

Figure 6. Top left: Original RGB image. Top right: NIR-G-B image (swapping NIR for R). Bottom left: R-NIR-B image. Bottom right: R-G-NIR image. Note that simply considering the NIR channel as a colour is not sufficient for a meaningful representation.

HSV

The question of which information channel is best suited to include near-IR info remains. The lightness channel V in HSV appears a suitable candidate for two reasons: V is de-coupled from colour information H and S, and the way colour is encoded is suitable for preserving the appearance of surfaces when the lightness information is changed. In this experiment, we basically swap the near-IR channel for the V channel of the visible image. We expect that by preserving the hue and saturation values of the original image, the colours will appear realistic and the image will be a combination of visible colours and near-IR intensity and spatial information. Figure 7 shows that while it is the case for most of the image, image regions that are significantly darker in the near-IR appear almost colourless in the combined image, a normal behaviour considering the inverted cone structure of HSV will push the dark values towards achromatism.

Figure 7. Top row: the V channel of the visible image (left) and the near-IR image (right). Bottom row: the original visible image (left) and the coloured NIR image (right). Note that while bright regions are generally well rendered, dark regions such as sky appear almost achromatic.

YCbCr

Another potential colour encoding to incorporate near-IR information is YCbCr. In contrast to HSV, while this space decouples luminance and chrominance, it does not necessarily preserve saturation. On the other hand, it should behave better than HSV when there is a significant decrease in luminance from the visible image to the near-IR one. Results from swapping the Y channel of the visible image for near-IR, Fig. 8, show that for the midtones, the outcome is similar to HSV's. Departure from these midtones, however, exhibit an inverse tendency to the HSV experiment. Here, one can see that the vegetation appears desaturated, while the sky regions are better rendered than previously, a logical result since perceived Saturation is the ratio of Chroma and Luminance and only the luminance is being modified here.

PCA and other Luminance-chrominance methods

When applied to colour images, principal component analysis generally results in a first principal component containing spatial information and another two related to colour content. Since we aim to replace the spatial (and lightness) content only,

---

[2] available at: www.csse.uwa.edu.au/ pk/research/matlabfns/

[3] ivrgwww.epfl.ch/supplementary_material/FS_CIC08/index.html

Figure 8. Top row: the Y channel of the visible image (left) and the near-IR image (right). Bottom row: the original visible image (left) and the coloured NIR image (right). Note that while dark regions are generally better rendered than with HSV, bright regions such as grass appear very desaturated.

we investigate the PCA approach to see if it provides an overall better representation than either YCbCr or HSV. To that effect, we first PCA the RGB image and normalise the near-IR image so that its range is comparable to the first principal component. In a second step, we replace the original principal component by the normalised NIR image and reconstruct an RGB image using the same transformation matrix. The results indicate two trends: on most images, the outcome of PCA is similar to the one of YCbCr (although with sometimes too saturated colours), and the results are greatly image-dependent. In fact, this method is unpractical since there are images for which the first principal component is not just in the direction of lightness variance but is also tainted with chromatic information.

Other possible luminance-chrominance colour encodings, such as Lab, Luv, or IPT, have also been investigated, but the resulting images are very similar the ones from YCbCr, so detailed results are not mentioned here.

Frequency methods

As near-IR data can be seen as spatial information, we want to investigate whether standard image processing techniques can be used to enhance the visual image with near-IR information. To that effect, we look at image frequencies and the phase of the images' Fourier transform.

The first approach consists in separating the visible and near-IR images according to their frequencies. A low-pass, gaussian, filter is applied to the images to obtain their low frequencies. The high-frequency images are calculated as being the original image minus its low-frequencies. Once the high- and low-frequency images have been found for both IR and visible images, we "swap" the high frequency component of the near-IR image with the ones of the RGB image. The idea is that since edges are high-frequencies, using the near-IR ones should enhance the original image's contrast. Unfortunately, results, such as the ones shown in Fig. 9 (middle image), indicate that this enhancement is either too small (most regions are unaffected) or too strong (ringing artefacts).

The second approach is to use wavelet transforms. We decompose the luminance of the visible and near-IR images using a symmetrical 8-band wavelet filter and 8 levels of wavelet decomposition. The wavelet coefficients of each decomposition are then compared. We retain, for each pixel, the maximal coefficient between the near-IR and the visible image for high frequencies, and we average these coefficients for low-frequency parts of the image. The new coefficients are then taken back in the spatial domain and the enhanced image is obtained. Results from this approach can be seen in Fig. 9.

Blending

Simply replacing channels might be a too naive approach. With the frequency methods, we have already investigated the possibility of retaining only one part of the near-IR data rather than its totality. This blending approach can take two forms; one mixes the Y, or V channel of the visible image with the near-IR image, effectively attenuating the influence of the near-IR contrast and saturation changes, or one can selectively replace the Y or V value of the visible channel by the near-IR, depending on whether there is an increase or decrease in the signal's intensity.

The first method's results, Fig. 10, show that while intermediate images exhibit less contrast and change, their overall visual quality is not greatly improved.

The second blending method is slightly more complex, but yields much better results. The RGB image is transformed to HSV and its V channel is compared to the near-IR channel. We then create a binary difference mask, depending on whether the visual V channel is higher or lower than the near-IR channel at a given pixel. If the near-IR indicate an increase, the V values of the corresponding pixels are replaced by the near-IR response at that location. For the other pixels, we go to YCbCr and replace their Y value by the near-IR response. This "in-between" transform in fact allows us to use the best behaved colour encoding for each transition, as shown in Fig. 11.

Psychophysical Experiment

To learn observers' perspective about our modification of the colour images, we have run an experiment with 10 observers and 8 images. We first devised nine possible modifications: four in YCbCr colour encoding where the Y channel is composed of 25%, 50%, 75% and 100% of the near-IR channel; four, with the same percentages in the HSV colour encoding; finally, the ninth correction is the "HSV or YCbCr" described above.

The experiment had two parts: in the first one, observers were asked to order the nine corrections in order of preference, and this for the eight images. The two preferred corrections were then used in a simultaneous test with the original RGB image, where observers had to indicate the image that was the most pleasing.

Results-wise, the overwhelmingly preferred correction (for 7 of the 8 images) was the method using a different encoding depending on brightness differences. Regarding the preference of modified over original image, the results are more nuanced. Indeed, while we expected the results to be largely image-dependent, they happen to be more observer-dependent. With five observers preferring all the modified images, and four preferring all the original images.

Figure 9. Results obtained using frequency methods. The original image in the visible spectrum (top). The Gaussian filter, FFT method (middle); some ringing is visible. The wavelet coefficients fusion approach (bottom), which provide enhancements but with the presence of some artefacts.

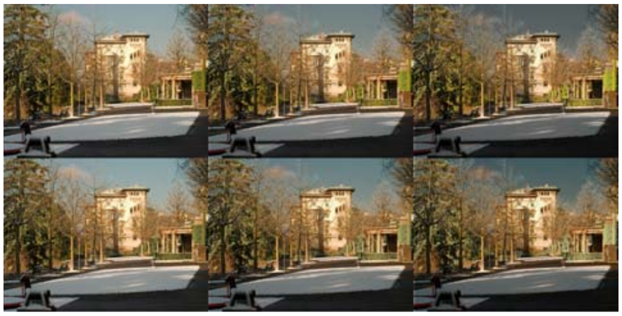

Figure 10. Top row: HSV-based rendering by blending the channels. From left to right, images where the V channel is composed of 25%, 50%, and 75% of the near-IR channel. The rest being provided by the original V channel. Bottom row: the same percentages applied to the Y channel of the YCbCr image.

Conclusions

Current digital cameras are intrinsically able to capture near-infrared images. In this paper, we have investigated physical phenomena that are of particular interest to digital photography and that can be enhanced by the use of near-IR images. Since most of these phenomena deal with contrast, brightness, and spatial information, we argue that near-IR images should not be treated as an additional colour channel but rather as a channel carrying luminance and spatial information.

Near-IR images are, however, grayscale images, so in order to incorporate them, we have experimented on various colour encodings and combinations thereof. A psychophysical evaluation was performed, and it was found that combining HSV and YCbCr colour encodings, depending on the direction of the modification gave the best results for image modification, however, when compared with the original RGB images, preference appears to be strongly user-dependent.

References

[1] S.Z. Li and R.F. Chu and S.C. Liao and L. Zhang. Illumination invariant face recognition using near-infrared images. *IEEE Trans. on Pattern Analysis and Machine Intelligence* pp. 627–639, 2007.

[2] T.M. Lilesand and R.W. Kiefer. *Remote sensing and image interpretation*. Wiley and Sons, 1994.

[3] D. Lu and E. Moran and M. Batistella. Linear mixture model applied to Amazonian vegetation classification *Remote Sensing of Environment* pp. 456-469(14), 2003.

[4] N. Milton and B. Eisworth and C. Ager. Effect of phosphorus deficiency on the spectral reflectance and morphology of soybean plants. *Rem. Sens. of Env.* pp. 121-127, 1991.

[5] D.A. Burns and E.W. Ciurczak. *Handbook of near-infrared analysis*. NIR Publications, 1999.

[6] M.D. Blue and S. Perkowitz. Reflectivity of common materials in the submillimeter region. *IEEE Trans. on microwave theory and techniques* pp. 491–493(6), 1977.

[7] M. Vilaseca and J. Pujol and M. Arjona and F.M. Martinez-Verdu. Color visualisation system for near-infrared multispectral images. *Journal of Imaging Science and Technology* pp. 246-255(3), 2005.

[8] M. Vilaseca and J. Pujol and M. Arjona and M. de Lasarte. Multispectral system for reflectance reconstruction in the near-infrared region. *Applied Optics* pp. 4241-4253(18), 2006.

[9] Kurt Nassau. *The Physics and Chemistry of Color: The Fifteen Causes of Color, 2nd Edition*. Wiley, 2001.

[10] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* pp. 91-110(2), 2004.

Figure 11. Left column: Original Images, Right column: Modified images using both HSV and YCbCr colourspace. The colours are better conditioned, while keeping the increased contrast and vividness of the near-IR images.