Memory Colors as Illuminant Predictors

Ted Cooper
Sony Media Processing Division
San Jose, California

Abstract

Segmenting an image can yield large objects whose generic color may be classified as memory colors. Typically sky, skin, foliage, and/or water are the dominant memory colors and can be found in the majority of personal photographs. An algorithm identifies potential memory color objects and uses their chromatic signature to predict the scene illuminant from a limited class of common light sources.

Overview

Because a DSC (Digital Still Camera) records an image as the product of the scene reflectance times the source illuminant times camera sensitivity at each pixel, the captured image represents a radiometric description of the scene. Late sunset images are very yellow in cast, whereas cloudy afternoon images are very blue in cast. This is an accurate representation of the chromatic information that a camera's CCD array receives. This is contrasted with the human visual system (HVS) that "discounts the illuminant" and remembers all the white scene objects as "white" even though the radiometric description given by the camera will be yellow-white for the sunset scene and blue-white for the cloudy day scene. To compensate for scene lighting conditions, it is necessary to predict the scene illuminant so that adjustments can be made to the raw camera data to make it acceptable and pleasing to the human eye.

A broad range of techniques is described in the literature to estimate the scene illuminant[1-9]. At one extreme exists the assumption that the mixture of scene reflectances is random and broadly distributed, so that a gray world approximation is valid. This permits the scene illuminant to be extracted from the pixel information by assuming that the average chromaticity for the entire ensemble of the image depends only on the illuminant's chromatic mixture. For a large number of typical scene conditions, this approxima-tion is valid, and the chromaticity of the scene is the chromaticity of the illuminant. But in a number of cases, dominant colored objects exist in the scene, and the gray world approximation fails. Techniques using statistical modeling of possible colors available under various standard illuminants have been employed. Their success depends largely on how chromatically distributed are the objects in the image. If bright blue and green objects are present, then these statistical models can reasonably maximize the probability that only a few types of illuminants would have that mixture possible. If two or more illuminants could yield the same type of chromatic results, then the statistics relating to the most probable mixture of colored objects will result in the final selection. If a scene composed of soft yellow-orange pastels were to be captured, then almost any illuminant from daylight to fluorescent could generate the chromatic mixtures of objects. The probabilistic methods would be at a severe disadvantage for such mildly colored images. Other illumination estimation methods rely upon hypothesis testing of standard illuminants incorporating basis vectors for both the illuminants and standard reflectance patches which might be found in the scene.

All of these methods depend upon a random or distributed mixture of colored objects in the scene to give a probabilistic selection of a given illuminant. This paper deals with an illuminant detection method that presupposes a specific type of object in the scene – not a random collection. This method assumes that at least one of a very limited set of memory color has been imaged. Having found such an object, an algorithm searches for a chromatic color signature for that object which is consistent with a specific illuminant.

Studies by Kodak in the photofinishing business have shown that 80% of all consumer film images contain people, plants, trees, and/or sky. Most pictures involve capturing people or landscapes. Almost all outdoor images will capture some form of foliage, in the various forms of trees, grass, plants, flower leaves, or bushes. With the exception of indoor product images (like brochure photographs of watches and jewelry), there is a very high probability that skin, sky, or foliage will be present.

In a previous paper, the use of memory colors was added to an algorithm for White Balancing (WB) as a way to provide robustness when near-neutral objects were not present[10]. The use of memory colors was a "safety net" for those rare instances when neutrals were not present but when sky, skin, and foliage might be. The limited results showed that memory color could predict a limited set of possible scene illuminants. This paper extends that work by adding a new memory color – water – and attempts to provide a statistical measure of how accurately the use memory colors can correctly impute the scene illuminant.

Methodology

The algorithm first employs a segmentation operation[1] to find all the large objects in the image (the methods are described in the previous paper). The size of the object is restricted by demanding that potential objects be 8-way contiguous with their neighboring pixels, and that it represent at least 0.5% of the total pixels in the scene. The algorithm permits heuristic growth of chromatic and luminance values of potential objects to find their full extent knowing that large brightness and minor chromaticity changes are expected because contoured shapes. In the previous paper, these large objects were then constrained to be near-neutral in chromaticity, and so a discrimination took place to remove any strongly colored objects. Of the four classes of memory colors considered in this paper, only skin comes near to being considered a near neutral object. The search for sky, foliage, and water typically requires fairly colorful chromaticity and so the algorithm had to be "tuned" to look for specific regions of color in order to identify memory color objects.

A choice of color space was necessary in order to take the raw CCD data and compose an image containing objects. RGB, L*a*b*, and HSL were all tested to see if one space had more robustness or speed over the others. The results indicated that camera RGB offered the most advantages. Since the raw RGB data has not been white balanced, it is very difficult to impute a white point for the L*a*b* space which is meaningful for this experiment. HSL suffers from the same problem to compute a luminance, L. In camera RGB space, an equal energy white (EEW) calibration of the raw data was performed so that color regions for red, green, and blue look realistic as a first order approximation. This pre-whitening step helps offset the large green multiplication factor present in most CCD arrays. Even with this EEW approximation, a fluorescent image will still exhibit a slight green cast. Similarly, an incandescent image will have a noticeable yellow cast. In either case, the green leaves of a tree will still fall in the green portion of the color spectrum and apples fall into the red.

The fundamental EEW calibration for the cameras employed in our measurements is close to D50 illumination, and typically had a Correlated Color Temperature (CCT) near 4900° Kelvin. Figures 1a and 2a show an indoor and outdoor scene with the EEW rendering from the raw data. Both images have foliage content and a MacBeth color checker chart to help test the accuracy of the memory color algorithm. With a LightSpex spectroradiometer, the two scenes had measured CCT values of 3565° K and 6637° K, respectively.

Figures 3a and 4a show the results for two outdoor images that have large bodies of water present. Because swimming pool water has such a tremendously different color spectrum than that of the ocean harbor scene, the segmentation algorithm had to be specially tuned to look for several different classes of water. Experimentation showed that the following classes were the most successful: azure water (pools, Caribbean coastlines, etc), and sea green water (deep ocean water and deep lakes). No special segmentation classifications were required for sky, skin, or foliage – aside from the considerations mentioned in the results section.

Once segmentation is completed, the algorithm needs to classify the large objects according to the general ranges of memory colors: sky, skin, foliage, water1 (azure), and water2 (sea). When at least one object is found, the algorithm sets confidence levels on the object(s) based on spatial location (are they horizontally or vertically oriented), and closeness to the central color region of the memory color class.

The algorithm then proceeds to verify the consistency of the illuminant predictions from the object(s). If two or more objects are found, do they all give the same illuminant type (daylight versus fluorescent versus incandescent)? If there is an inconsistency, the algorithm looks for the objects of the largest pixel extent. While the algorithm is still under development, most of the test cases involving known memory color objects have been correctly identified, and yield the correct illuminant type. Figures 1-4b show the rendering of the raw data based upon the imputed illuminant from the memory color objects that were found.

Results

Most memory colors are biased towards outdoor environments where daylight is the dominant illuminant. With the exception of skin and indoor plants, all of the natural occurrences of sky, skin, foliage, and large bodies of water are found outdoors. In capturing images for this research, special attention had to be paid in the selection of scenes so that a significant number of indoor light sources (fluorescent and incandescent) were present. This mostly was satisfied by human portrait scenes and capturing of large-leafed indoor plants.

Having selected camera RGB as the color space of choice, next the mathematical form to represent chromaticity was needed. Previous research with film drum scanners and flatbed scanners has shown that R/G and B/G give a meaningful chromatic representation[11]. Since green filters typically transmit more luminance information that blue or red, the R/G and B/G ratios are always less than one for all practical camera and scanner filter arrays. With this historical background, the R/G and B/G chromaticity measure was selected. After completion of the initial research, it was decided to reprocess all the data in a G/B and R/B chromaticity space to determine if other color representations would yield conflicting results. G/B and R/B ratios resulted in the same illuminant prediction. However, the noise level increased because many of the memory colors possess small blue components, and hence the small denominator often increased the noise level. The other color space of G/R and B/R was processed, but it provided even more noise enhancement for the water samples since very little red is present.

Figures 5 through 8 graph the results for images captured under various illuminants with memory color objects present. The R/G and B/G ratios of these segmented objects are compared with the equivalent ratios for the set of illuminants: D65, D50, CW fluorescent, Horizon, U30 fluorescent, type-A incandescent, and a 3200 Kelvin studio light. In all cases, the correct type of illuminant (daylight, fluorescent, incandescent) was chosen. In some instances, D50 was closer to the result but D65 more closely matched the measured CCT for the scene. However, the minor difference of daylight illuminants would not significantly alter the white balance of the scene rendering compared to the white balance of the measured illuminant. The P2, P3, and P4 notations in the figures refer to the MacBeth patches for Light Skin, Blue Sky, and Foliage, respectively. The large symbols in these figures are reserved from the MacBeth patch under different illuminations to act as a reference. The small symbols indicate actual measurements from various images, including one P2 patch in the scene.

Sky, skin, and foliage gave very consistent results for outdoor scenes. More error was encountered for indoor scenes with fluorescent illuminants, but this only accounted for a single error out of 25 test scenes. The largest error resulted from the use of water as a memory color. Even with the sub-division of water into a dark blue and light blue components, there were 7 cases out of 14 where water classification of specific illuminant was incorrect. But even for water objects, only 1 case was identified where the wrong type of illuminant (daylight, fluorescent, incandescent) was selected. As can be seen in Figure 8, the dark sea objects span a large range of R/G. The P6 patch (Bluish Green) is the closest MacBeth patch to match dark sea, and is used for illustrative purposes only. The P6 patch is a much better match for azure water (like swimming pools). Extreme care must be taken when using water memory color objects not to include specular type reflections in the segmentation area. When this happens, the water takes on the dominant color of the reflected object. This can be avoided by restricting the segmentation routine to ignore “bright islands” within the water region. Typically this is handled by brightness restrictions within the segmentation routine.

Figures 9 and 10 show the large distribution of R/G and B/G ratios for two types of water, pool and ocean respectively. The large symbols in these two figures are reserved to indicate water patches for which medium to strong surface reflections are included. Note how the addition of surface reflections in Figure 9 significantly extends the R/G and B/G range. Some pool water reflections showed enough red content to belong to a horizon illuminant, even when the scene was captured in mid-afternoon with a CCT for D68. In Figure 10, the sea green ocean patches also occupy a large span in the R/G direction. However, they do not show any significant changes when areas with surface reflections are included. This results from the fact that ocean water contains a large amount of algae and suspended solids. There is a strong chromatic cast to the light reflected both on and below the surface of the water.

The results of this research indicate that sky, skin, and foliage objects retain good signatures of the scene illuminant. Dark skin proved unreliable under certain lighting conditions (A and Horizon) and reliable under others. Typically with low luminance levels, the low reflectance produced very erratic R/G and B/G values. The use of water as a memory color is very problematic. If a green sea image has half the horizon filed with ocean, the algorithm will indicate the daylight source with no problem. When the size of the “water” object becomes less than 15% of the image, the potential water object can very often be a blue wall or a blue-green couch. More research needs to be conducted on water scenes to see if additional constraints on the objects surrounding the potential water object can increase the confidence level in predicting a water memory color. Constraints such as requiring earth colors or sand to be adjacent to the potential water object could significantly reduce false positive identifications.

The memory colors of sky, skin, and foliage reliably return the source illuminant type (daylight, fluorescent, incandescent). Water for dark sea green works well when large areas of the image contain the memory color object.

Conclusions

The dominant memory colors of sky, skin, and foliage worked successfully 96% of the time. Even the single failure in our 25 test cases occurred because the foliage was very dried out and dying. Sky gives the most accurate illuminant because it is always involved in daylight cases. The only problem that can be encountered with sky segmentation is including too much cloud material along with the blue sky. However, when this occurs, the clouds represent very strong near neutral objects. When clouds are present, the CCR white balance algorithm works correctly in all observed instances.

Skin is the best overall memory color because it is found both indoors and outdoors so all forms of illumination are encountered. All of our skin tests algorithm results yielded the correct illuminant. Some cases occurred where the measured CCT of the scene was between D50 and D65 so either answer was accepted as a correct answer. Skin appears to have the strongest discrimination between daylight and fluorescent sources. This is very important since a switch between daylight and fluorescent, where both might possess the same CCT value, has a dramatic effect on the final image reproduction. The 3200 studio light has roughly the same CCT as U30 fluorescent, but the resulting amplifier gains needed to provide good white balance are significantly different.

The worst performer as a memory color is water. The azure water is extremely bad because, for swimming pools and shallow bodies of clear water, little color reflection occurs since there is not enough material to cause chromatic dispersion. The addition of surface specular reflection can completely shift the chromatic content of the reflected light from light blue towards red, green, or white depending of the reflection content. Alternatively, sea green water, even with specular reflections present, provides a much better memory color results. However, the dispersion of R/G and B/G ratios for sea green water did not allow us to separate D65 from D50 or D50 from Horizon conditions. In our test images of boats in a marina, it was always possible to find an ocean reflection where a red or yellow boat would cause the resulting patch of ocean to have a Horizon or incandescent A illuminant signature, even though the real scene illuminant was D50. If a large enough region of the ocean can be captured (or a large variety of boats at a marina), then the global average of the sea green water does provide a good illuminant signature. However, this is really more of a gray world approximation for sea water than a success of memory colors.

In conclusion, the memory colors of sky, skin, and foliage are very good indicators of scene illuminant. Water has a limited success for ocean water but probably is too unreliable to be trusted as anything else than supporting evidence that a given illuminant has been detected by other means. Future research needs to investigate how to construct simplistic segmentation routines that give a high confidence that sky, skin, or foliage are present in an arbitrary image.

References

1. T. Cooper, I. Tastl, and B.Tao. “Novel approach to color cast detection and removal in digital images”. Proc. of SPIE on Color Imaging: Device-Independent Color, Color Hardcopy, and Graphic Arts V. Vol. 3936, pp.167-177, 2000.
2. Hsien-Che Lee and R. Goodwin, “Colors as Seen by Humans and Machines”, Proc. of the IS&T 1994 Conference Proceedings, P.401.
3. R.W.G. Hunt, “A model of colour vision for predicting colour appearance in various viewing conditions”, Color Res. Appl. 12(6) pp 297-314 1987.
4. K. Takemura, “Photographic prints from digital still cameras”, Proc. of SPIE on Color Imaging: Device-Independent Color, Color Hardcopy and Graphic Arts IV SPIE Vol.3648, pp 158-166 1999.
5. G. Finlayson, P. Hubel and S. Hordley. “Color by correlation”. Proceedings of the Fifth IS&T/SID Color Imaging Conference. 1997.
6. B. Funt, V. Cardei and K. Barnard. “Learning color constancy”. Proceedings of the Fourth IS&T/SID Color Imaging Conference. 1996.
7. H. Trusell and M. Vrhel. “Estimation of illumination for color correction”. Proc. International Conf. Accous. Speech Signal Processing. 1991.
8. L. Maloney and B. Wandell. “Color constancy: a method for recovering surface spectral reflectance”. J. Opt. Soc. Am. A Vol. 23, #1. 1986.
9. B. Tao and I.Tastl. “Illumination Estimation using Linear Decomposition and Constrained Optimization”. IS&T/SPIE Color Imaging Conference. Jan. 2000 (to be published).
10. T. Cooper, “Color Segmentation as an Aid to White Balancing for Digital Still Cameras”, Proc. of SPIE on Color Imaging: Device-Independent Color, Color Hardcopy and Graphic Arts VI SPIE Vol.4300, pp 164-171 2001.
11. Kobus Barnard,”Modeling Scene Illumination Coulour for Computer Vision and Image Reproduction: A survey of computational appoaches”. A Ph.D. thesis which can be found on the Web at: www.cws.berkeley.edu/~kobus/research/projects/depth_exam/depth_exam.html p12.

Biography

Dr. Cooper graduated from M.I.T in 1974 in Physics. He has worked in a number of image processing fields including medical and satellite imaging. For 12 years he as performed research in graphic arts community in color printing technology. For the past three years he has worked at Sony in the area of digital cameras and color enhancement.

Figure 1. a) Indoor scene with Plant; b) after U30 fluorescent rending

Figure 2. a) Outdoor scene with grass; b)after D65 rending applied

Figure 3. a) Swimming pool in afternoon sun; b)after D65 rendering applied

Figure 4. a) Boats in marina in afternoon sun b)after D65 rendering applied

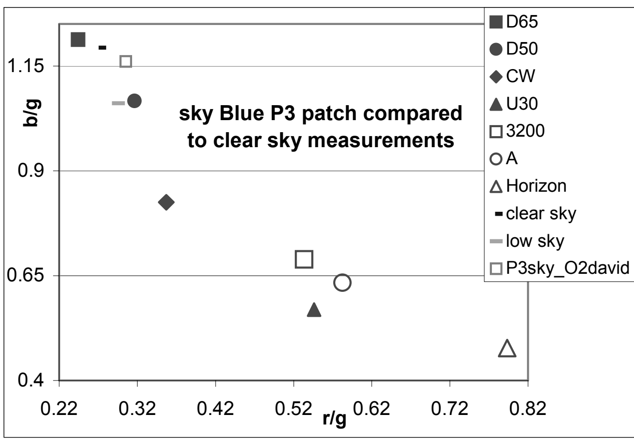

Figure 5. r/g versus b/g ratios for sky

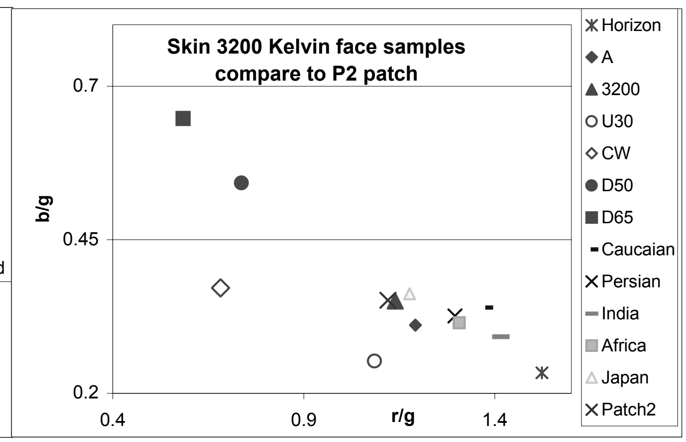

Figure 6. r/g versus b/g ratios for skin

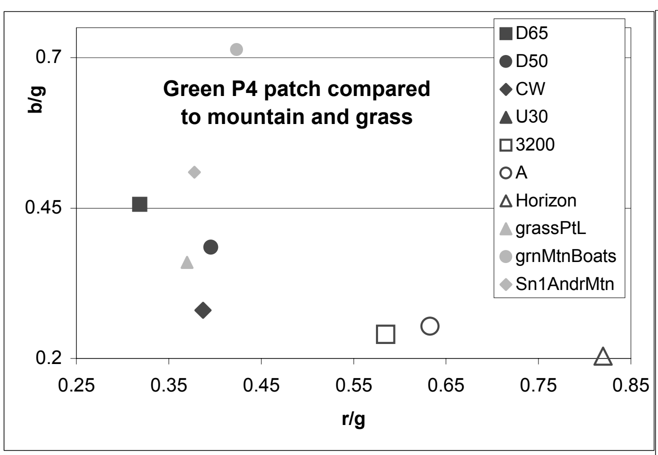

Figure 7. r/g versus b/g ratios for foliage

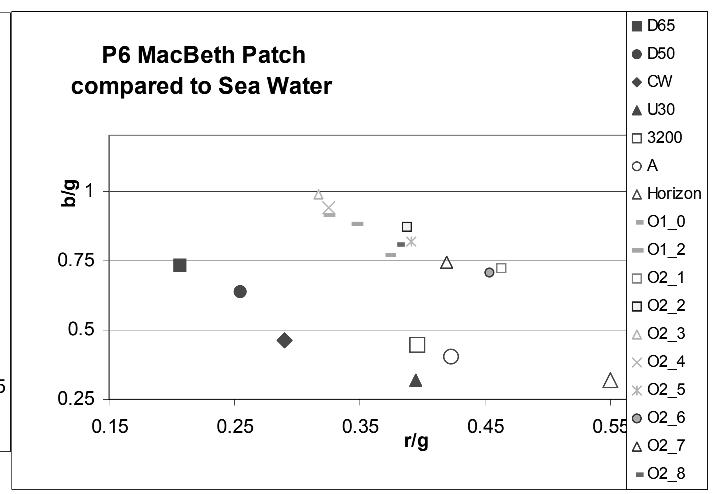

Figure 8. r/g versus b/g ratios for sea water

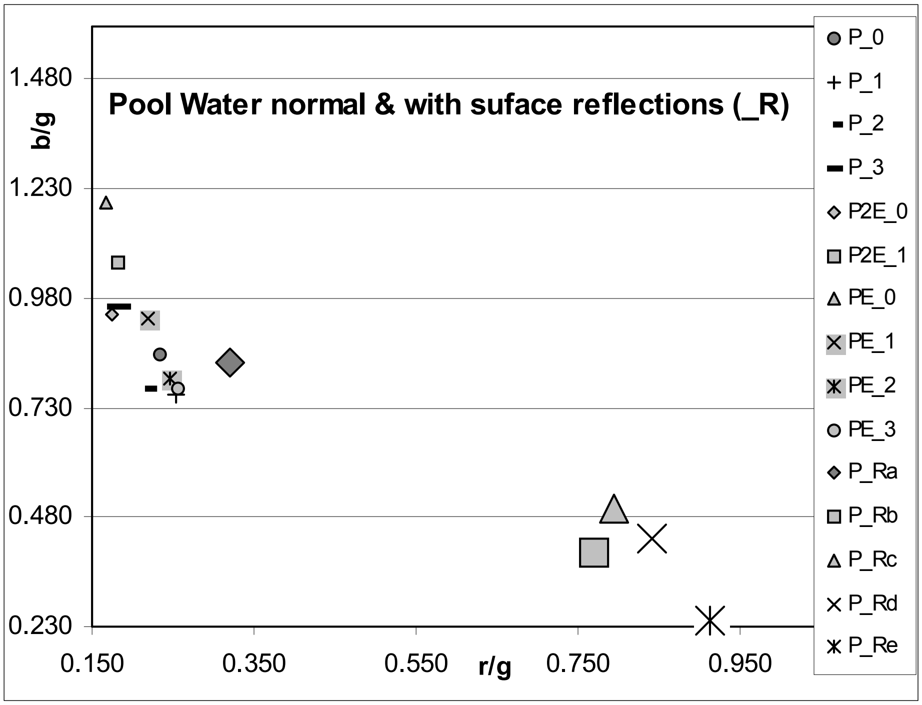

Figure 9. r/g versus b/g ratios for Pool Water with and without surface reflections

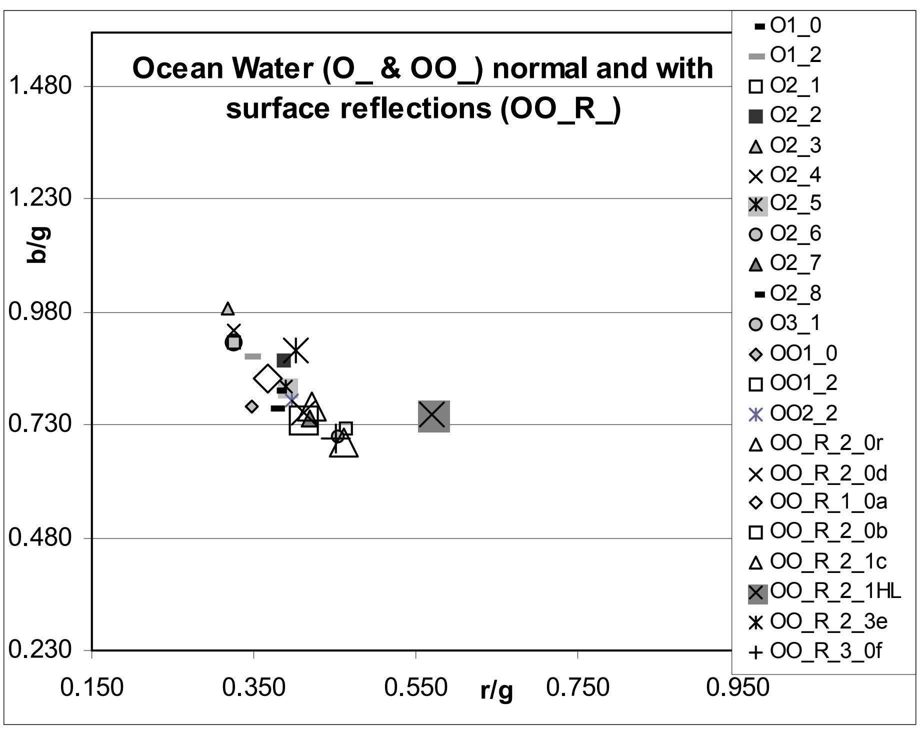

Figure 10. r/g versus b/g ratios for Ocean water with and without surface reflections