

Colour reconstruction of underwater images

Julian Hoth^{*1,2} and Wojciech Kowalczyk¹

¹ University of Duisburg-Essen, Lotharstr. 1, 47057 Duisburg, Germany

² German Aerospace Center (DLR), Institute of Communications and Navigation, Department of Nautical Systems, Kalkhorstweg 53, 17235 Neustrelitz, Germany

Received 2 March 2016; Accepted 11 July 2016

Abstract

Objects look very different in the underwater environment compared to their appearance in sunlight. Images with correct colouring simplify the detection of underwater objects and may permit the use of visual simultaneous localisation and mapping (SLAM) algorithms developed for land-based robots underwater. Hence, image processing is required. Current algorithms focus on the colour reconstruction of scenery at diving depth where different colours can still be distinguished, but this is not possible at greater depth. This study investigates whether machine learning can be used to transform image data. First, laboratory tests are performed using a special light source imitating underwater lighting conditions, showing that the k -nearest neighbour method and support vector machines yield excellent results. Based on these results, an experimental verification is performed under severe conditions in the murky water of a diving basin. It shows that the k -nearest neighbour method gives very good results for short distances between the object and the camera, as well as for small water depths in the red channel. For longer distances, deeper water and the other colour channels, support vector machines are the best choice for the reconstruction of the colour as seen under white light from the underwater images.

Keywords: Colour reconstruction, underwater camera, unmanned underwater vehicles, marine robotics

1. Introduction

Objects look very different in the underwater environment compared to their appearance in sunlight. The main reason is that the penetration of light through seawater is highly dependent on the wavelength of the light (Åhlen, 2005). Suspended particles in the water can further decrease the overall quality of underwater images (Bazeille et al., 2006; Celebi and Ertürk, 2012).

High quality images with correct colouring simplify the detection of underwater objects and may allow the use of visual simultaneous localisation and mapping (SLAM) algorithms developed for land-based robots underwater. Hence, image processing is required to obtain images of high quality and correct colouring. Over the last decade, significant progress has been made in this direction.

Current algorithms focus on the colour reconstruction of scenery at diving depth (Iqbal et al., 2007; Celebi and Ertürk, 2012). Therefore, a significant part of sunlight is still present and different colours can still be distinguished, although they may be tainted due to light being filtered through seawater. The algorithms are often based on a simple relationship between colours under different lighting conditions. Typical image models are Beer's law (Åhlen, 2005) and the Jaffe-McGlamery image model (Jaffe 1990; Lee et al., 2012), which are shown in Fig 1. Iqbal et al. (2007) use contrast stretching of the red-green-blue (RGB) image, and saturation and intensity stretching of the image converted to hue intensity saturation (HIS) colour space in order to correct the colour. However, stretching requires a certain minimum amount of intensity information retained in the image. This is only true for small depths and the examples given in the paper fulfil this requirement. Other approaches require known geometric relations between the camera and the object (Lee et al., 2012) or require manual work (Chen et al., 2014).

Unfortunately, at greater depth the filtering is much stronger, such that different colours are strongly tainted and can no longer be distinguished (Fig 2). Hence, simple image models are no longer applicable. To solve this, the main factor of interest for the influence on colour is the penetration depth (d_p) which is a strong function of the wavelength (λ) (Fig 3).

* Contact author. Email address: julian.hoth@dlr.de



Fig 1: Pairs of images before (left) and after correction (right) from Bazeille et al. (2006)



Fig 2: An object under normal lighting conditions (left) and under underwater lighting conditions (right)

This study investigates whether machine learning can be used to transform image data. This should work for extreme conditions as shown for example in Fig 2. To this end images of objects of different colour are obtained under underwater and regular lighting conditions.

2. Machine learning

This section briefly outlines the machine learning procedures that were used in this study.

2.1. Artificial neural network

Artificial neural networks (ANNs) are information-processing algorithms that are modelled after

the way human brains work (Aleksander and Morton, 1995). The first formulations of this method were already made in 1943 by McCulloch and Pits.

Artificial neural networks consist of strongly interconnected nodes called neurons. These neurons are organised into layers: one input layer, one output layer, and one or more hidden layers in-between. The number of nodes in the input and output layer is given due to the nature of the data being processed. In this study, there are three nodes in each of them to represent the colour channels of a 24 bit pixel. The number of nodes in the hidden layers as well as the number of hidden layers is variable (Ferreira, 1996).

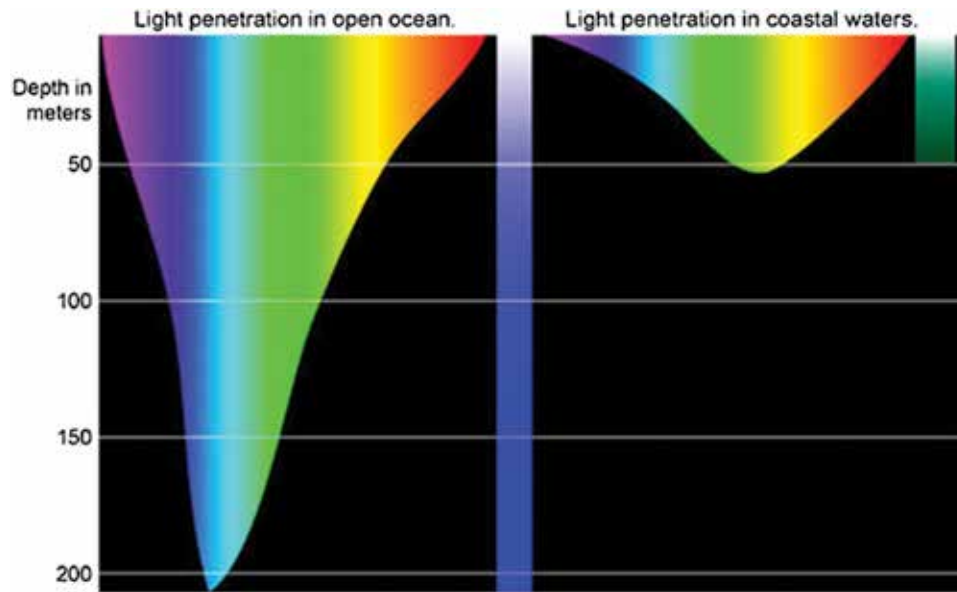


Fig 3: Light penetration in ocean water and coastal waters (Carothers 2016)

Teaching neural networks is done by forming or removing connections, changing the weights and the threshold of neurons, and adding or removing neurons so that errors are minimised for a given validation set. However, in many cases the overall shape of the neural network is fixed and only the weights and thresholds are used for learning (Ferreira, 1996).

A common method for training is the back-propagation algorithm, which is also used in this study. Hereby, the error derivative of the weights – i.e. the change of error depending on the change of weight – is determined by starting with the total errors at the output layer and the moving through the network towards the input layer (Rojas, 1996). This is done repeatedly until either the rate of change of the errors or the error derivatives becomes sufficiently reduced.

In this study, up to 250 learning cycles are used with a learning rate of 0.1 and a momentum of 0.2 for a given neuron configuration. The number of neurons was varied between 3 and 8 per layer with the number of layer between 1 and 4.

2.2. *k*-nearest neighbour

The *k*-nearest neighbour (KNN) is a non-parametric method for classification and regression. The output (classification or property value) for a given input is obtained from known input-output relations where the inputs are similar and close (i.e. ‘in the neighbourhood of’) to the sample point in question. Classification is then done by majority voting and in case of a property value the average of the outputs of the neighbouring inputs is taken (Fix and Hodges, 1951). In this study, data points are situated in a 3D space (due to three colour channels) and can have 256 different values (classifications) for each channel.

The result depends on the choice of the neighbourhood – i.e. up to what distance or how many neighbours are taken into account, and the type of distance measure has some influence (Lewicki and Hill, 2005). In this study, Euclidean with weighted averaging is used. Teaching a KNN system is usually performed by finding the size of the neighbourhood *k* with the lowest error by cross-validation (Lewicki and Hill, 2005).

2.3. Support vector machines

Support vector machines (SMVs) are non-probabilistic linear classifiers. However, they can also be used for regression (Drucker et al., 1997). The fundamental idea is that data sets that belong to different classes are linearly separable by hyperplanes. If the data cannot be separated by linear hyperplanes it has to be mapped into a higher dimensional (embedding) space such that it becomes linearly separable (Cortes and Vapnik, 1995).

Of all those hyperplanes that separate the datasets, one hyperplane has to be chosen such that the margin separating different classes is a maximum. Those data points closest to the boundary and which are required to describe the hyperplane exactly are called support vectors. Hence, teaching a support vector machine is essentially an optimisation problem.

As for KNN, the support vector machine in this study starts in a 3D space and each data point is in one of 256 classes for each colour channel. Due to the likely non-linear separability the resulting machine has a much higher dimensionality.

SVMs with a Gaussian Kernel are considered to be the best choice for this study. Kernel parameter (σ) is varied between 0.1 and 5.0 to obtain the spread

with the lowest error. For the optimisation, the cost of misclassification is set to a moderate value of 1.0, the convergence tolerance is 0.0010, and the width of the ε -sensitive zone is set to $\varepsilon = 0.0010$.

2.4. Bayesian network

Bayesian networks (BN) are probabilistic graphical models based on Bayesian probability theory (Bayes 1763; Jaynes and Bretthorst, 2003). A Bayesian net describes how different states of a system represented as nodes of a graph are linked through probability, i.e. the net shows conditional interdependencies of variables via a directed acyclic graph (Ben-Gal, 2007).

Bayesian networks can be learned from available data. Various procedures are available for teaching, which are separated into two categories (Cooper and Herskovits, 1992; Heckerman et al., 1995):

- 1 Structure learning (structure is unknown); and
- 2 Parameter learning (structure is known).

Often, they are combined such that parameter learning is a sub-process of structure learning (score-and-search-based approach) (Friedman et al., 1997). In this study, hill climbing (Gómez et al., 2011) with an alpha parameter of 0.5 is used.

2.5. Multiple linear regression

Multiple linear regression (MLR) is similar to linear regression, but instead of having one independent variable as for linear regression, the output depends on two or more independent variables. In this study, there are three separate MLRs (one for each colour channel in the reconstructed image) with three independent variables (the three colour channels of the image under underwater lighting conditions). Learning a multiple linear regression machine is done via the method of least squares. The best linear unbiased estimator is found using the Gauss-Markov theorem.

3. Methods

3.1. Laboratory test

To obtain images under controlled lighting conditions, various coloured objects were illuminated by a special LED light source and regular white light. The special LED light source has a range of 450 nm to 570 nm and a mean wavelength of $\lambda = 498$ nm. The mean wavelength is very close to the wavelength for maximum penetration depth in open ocean, $\lambda_{max, ocean} \approx 480$ nm (see also Fig 3; Smith and Baker, 1981; Mobley, 2004). Therefore, the objects looked as if they were situated in deep ocean water and illuminated by a white light from a submarine

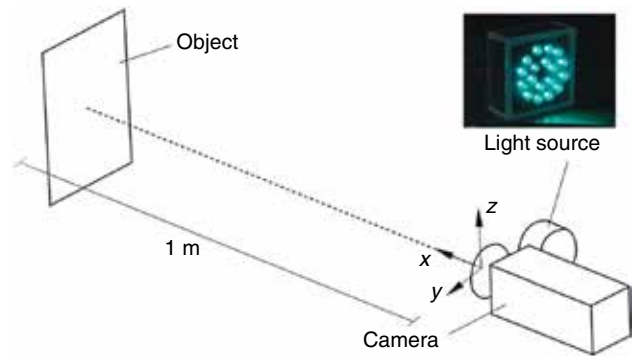


Fig 4: Laboratory setup for image acquisition

or underwater robot from some distance away. For the regular white light, a mercury-vapour lamp is used.

The coloured objects were then photographed several times under both white light and underwater lighting conditions. The camera used had an active pixel sensor with 16.1 mpx resolution. A sketch of the laboratory setup is shown in Fig 4.

The images taken were then processed as follows. First, feature-matching is done for each image set, where the same object is photographed once under white light and once under underwater lighting conditions so that the two images are exactly aligned. Therefore, every pixel in the image under underwater lighting conditions has a corresponding pixel in the image taken under white light. For feature-matching, the *matchFeatures* algorithm implemented in MATLAB[®] has been used. This method finds corresponding points of interest between pairs of images using local neighbourhoods and the Harris algorithm (MATLAB, 2016).

In a second step, the images were resized such that only the coloured objects could be seen in the pictures. This was necessary as the background was not illuminated. From every image set, 10% of the pixels were chosen at random and an $n \times 6$ matrix was built containing the colour channels (RGB) for two matching pixels in each row. The resulting matrix was then fed through statistical learning algorithms with or without pre-filters. The methods discussed in section 2 of this paper were chosen for testing.

The input data for the learning methods were also varied to test the influence of different pre-processing methods on the performance of the machines. Besides feeding the raw image data (raw) directly into the statistical learning algorithms, the following methods were used for smoothing the data:

- fast Fourier transform (FFT);
- convolution (Conv);
- moving average (MA);
- singular value decomposition (SVD).

The FFT smoother eliminates frequency noise by setting all frequencies to zero that are less than a threshold times the maximum distance of the data. The convolution smoother uses the Gaussian kernel. The size of the convolution filter was set so that the difference between the original and smoothed data is no more than 1% of the whole range of values. Moving averages was done with five consecutive entries, and the SVD smoother removes singular values that are below a chosen threshold.

3.2. Underwater experiment

The second stage was to test the applicability of learning algorithms for underwater image processing under more realistic conditions. This allows a more appropriate evaluation of the method before it is implemented into a real system. Hence, an experimental verification in water was warranted.

In order to be able to evaluate the results as best as possible and to automate the colour matching for the learning algorithms, the coloured objects from the laboratory test were replaced by colour patterns with known colours. Fig 5 shows one of these patterns. The patterns were such that they could easily be processed with standard pattern recognition algorithms to extract the colours from the images at the appropriate positions. One pattern consisted of 32 carefully chosen colour patches, including white, black, and several shades of grey as well as colours from the whole RGB range. The other colour patterns had random patches. The first pattern was intended for training of the machine learning algorithms. The remaining patterns were for testing the trained learning

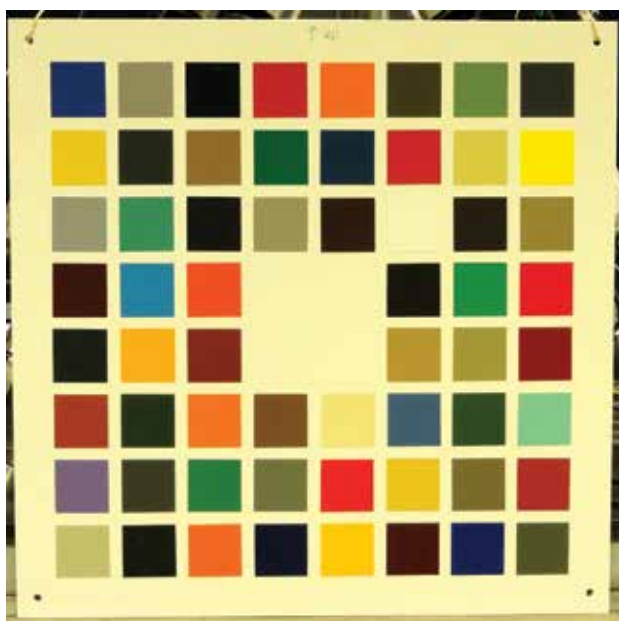


Fig 5: One of the colour patterns used for the experimental verification



Fig 6: Diving basin at TauchRevier Gasometer

machine on various colours that were not part of the training process.

The colour patterns were first photographed in the laboratory under white light at distances of 1 m, 5 m and 8 m. This was done for various camera and zoom settings to determine the best conditions. The camera used had an active pixel sensor with 16.1 mpx resolution.

The boards were then brought to Europe's largest indoor diving basin at TauchRevier Gasometer in Duisburg, Germany. The diving basin is circular with a diameter of 45 m and a water depth of 13 m. It provides conditions that are as close to open water as possible (Fig 6), and it is therefore rated as open water for divers. Furthermore, the lighting conditions are such that greater water depths are simulated compared to the actual water depth achieved. Also, the optical properties of the water are similar to coastal water and the water is very murky. Therefore, the penetration depth of light was expected to be very low in general, with its maximum in the green colour region.

The boards were taken to depths of 4 m, 8 m and 12 m and again photographed at distances of 1 m, 5 m and 8 m using the same camera that is used in the laboratory (Fig 7). Under white light, this is done for various camera and zoom settings. The resulting images were processed as for the laboratory test.

4. Results and discussion

4.1. Laboratory test

Table 1 shows the root-mean-square errors (RMSEs) for all combinations of the five machine learning methods and the five different input data for all colour channels of 24 bit images. Every channel has 256 possible pixel values for the laboratory test. As it is only the RMSEs, the actual deviation of a pixel colour from the expected value can differ

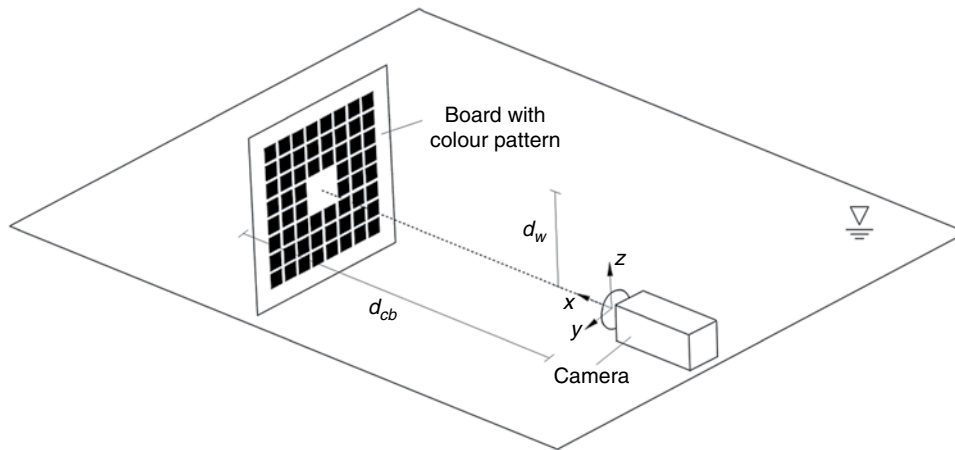


Fig 7: Experimental setup (d_{cb} : distance camera/board, d_w : water depth)

Table 1: RMSEs for different learning methods and pre-processors in laboratory setup

RSME	ANN	KNN	SVM	BN	MLR
Raw	57.06	30.46	12.35	60.40	43.89
FFT	56.84	9.32	11.38	60.40	43.89
Conv	54.06	10.70	11.76	60.40	43.90
MA	62.30	80.84	48.12	62.16	43.91
SVD	56.93	30.46	11.35	60.40	43.89

significantly. The RMSEs for the red channel are slightly higher while the ones for the blue channel are slightly lower than the values given in the table.

Bayesian networks and ANNs show a large RMSE independent of the input data of well above 50 px. MLRs also show a significant RMSE of around 44 px for all types of input data, although they are lower than for ANN and Bayesian networks. For KNNs and SVMs the performance depends on the input data. Both show a very large RMSEs when the image data were smoothed with moving averages. In fact, KNNs with moving averages has the largest error (almost 81 px) over all combinations, while KNNs with raw data or data decomposed with singular value decomposition shows a medium RMSE around 30 px. Much lower RMSEs (about 12 px) are observed for SVM using any type of input data except data smoothed with moving averages. The lowest RMSEs are obtained for KNN combined with convoluted data and FFT. The low performance of Bayesian networks can be expected as image data are not probabilistic. The high RMSE for MLR was also anticipated. As stated earlier, the assumption of a linear dependency between the colours as seen underwater and under sunlight does not hold beyond small depths.

Smoothing the input data reduces the performance of the learning machines. Especially for KNN and SVM where the RSME is quite low, the drop in performance is significant. It can therefore be deduced that small changes of the pixel values

have a strong influence on the result. These small changes are removed by smoothing and hence the input data no longer represent the original image properly.

Fig 8 shows how the approach in this study works for some sample objects when using KNN with FFT smoother in the laboratory, which is the combination with the lowest error. For every object, there is a set of three images: (a) the object under white light; (b) the same object illuminated by the special light source (see Fig 4); and (c) the reconstruction done by the algorithm using the middle image as an input.

As Fig 8 shows, the colours of the object under white light are reproduced well. However, there are also some situations where the reconstruction is not optimal. Problems occur at mid left and at the top edge in Fig 8(c), where strong red colours are produced, because red light is strongly influenced by the lighting conditions (see Fig 3). It can be observed that red objects look almost the same as black and dark grey objects. Hence, regions which appear in these colours in the underwater image have to be mapped both to red and to black or dark grey. It is clear that this does not work in every situation.

It can also be seen in Fig 8 that due to the strong and focused light source as well as the smooth surface of the sample objects, reflections occur in some regions of the image. The scenery is not uniformly lighted as for the object under white light. In these parts the colour reconstruction is not as good as in the rest of the image. Reflections change the relation between the colours under white light and under underwater lighting conditions.

4.2. Underwater experiment

Due to the conditions onsite, it was not possible to obtain images of every colour pattern for every water depth and camera distance. In addition, not all combinations of distance and depth gave feasible results. Table 2 shows an overview of the performance.

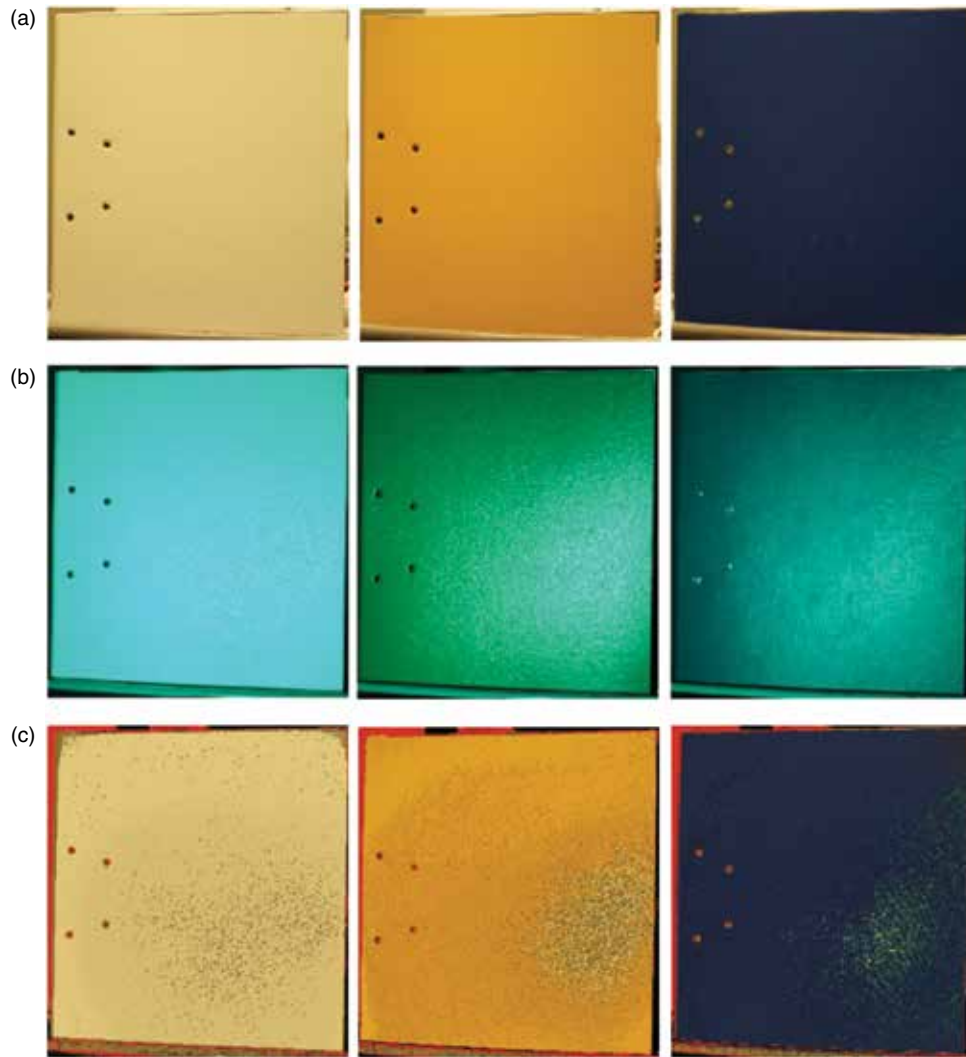


Fig 8: (a) Object under normal lighting conditions, (b) under underwater lighting conditions and (c) reconstruction from learning machine

Table 2: Performance of the underwater tests

	Camera distance [m]			
	1	5	8	
Water depth [m]	4	All patterns	All patterns	Only one pattern
	8	All patterns except one	All patterns except one	No images
	12	Only two patterns	No feasible images	No images

No feasible images could be obtained at a water depth of 12 m and a camera distance of 5 m. Fig 9(b) shows an image taken of pattern 2 under these conditions. As can be seen, only the measurement cord is visible in the left part of the image; the light cannot penetrate the water to the colour pattern itself. The same is true for a camera distance of 8 m and water depths of 8 m and 12 m. Hence, no images were obtained for these conditions. Fig 9 also shows some further sample images for various camera distances and water depths. These give a good indication of the visibility conditions onsite.

As there were numerous colour patches under different conditions, only the worst case (water depth 12 m, camera distance 1 m) for selected colour patches is discussed here in more detail. The observations made can also be applied to the other cases.

Table 3 shows the RMSE for all combinations of the five machine learning methods and the five different input data for all colour channels of 24 bit images, i.e. every channel has 256 possible pixel values, for a water depth of 12 m and a camera distance of 1 m. The RMSEs for the red channel are 10 px to 15 px higher, while the RMSEs for the green channel and the blue channel are 5 px to 10 px lower than the values given in the table.

The results show a similar structure compared to the results for the laboratory experiment. Bayesian networks and MLR show a better performance in the experiment than in the laboratory test. For ANNs, KNNs and SVMs, the errors are increased. For each learning algorithm, the errors for the different input methods are very similar except for KNNs where a strong variation between 49 px and

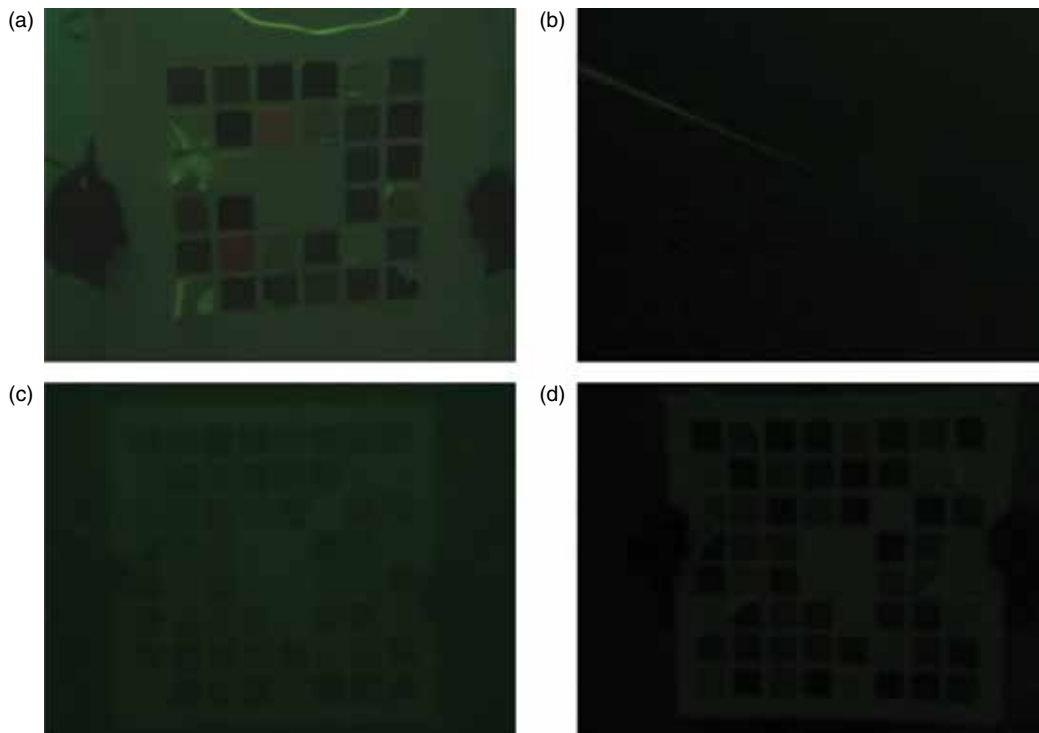


Fig 9: Sample images from the underwater experiment for (a) 4 m depth and 1 m distance, (b) 12 m depth and 5 m distance, (c) 8 m depth and 5 m distance and (d) 12 m depth and 1 m distance

Table 3: RMSEs for different learning methods and pre-processors at 12 m water depth and 1 m camera distance

RSME	ANN	KNN	SVM	BN	MLR
Raw	66.52	61.13	31.04	45.50	37.19
FFT	66.55	52.69	31.73	44.78	36.85
Conv	68.03	54.08	31.91	44.52	36.86
MA	68.35	70.41	48.41	44.05	36.89
SVD	66.15	49.98	32.81	45.50	36.85

71 px can be observed. Moving averages also decrease the performance of SVMs significantly. The lowest RMSEs are obtained for SVMs combined with raw data, convoluted data and FFT.

Table 4 summarises the results for the different camera distances and water depths used in the experiment. It shows which learning machines give the best results separately for the three colour channels. SVMs become dominant the further away the camera is from the object and with increasing water depth. KNN is only used for small camera distances and small water depths, and is of real interest only for the red channel alone. For the green channel, KNN gives similar results to SVMs at a camera distance of 1 m and a water depth of 4 m, but in no other case. For the blue channel, only SVMs are used.

It may therefore be concluded that SVMs are generally the better choice for the task of obtaining the colour of objects under white light from the underwater images, while KNN should be used only at close distances, though this is not of interest for the task of underwater navigation and object recognition.







































Table 4: Learning machines used

		Camera distance [m]		
		1	5	8
Water depth [m]	4	R: KNN	R: KNN	R: SVM
		G: KNN/SVM	G: SVM	G: SVM
		B: SVM	B: SVM	B: SVM
	8	R: KNN	R: SVM	
		G: SVM	G: SVM	
		B: SMV	B: SVM	
	12	R: SVM		
		G: SVM		
		B: SVM		

Having analysed the channels separately, it is also necessary to discuss the colours produced by the algorithms from the underwater images. Table 5 shows the results for several colour patches when using SVMs with raw data input, which is the combination with the lowest error. The left column shows colour patches as they appear in the laboratory under white light. In the middle column, the same patches are given as they look in the underwater environment in the water tank at 12 m water depth and 1 m camera distance. The right column shows the reconstruction from the algorithms.

The first general observation that can be made is that there is no perceptible difference between the colour patches in the underwater environment, so there is very little information contained in the images that can be used for the reconstruction. In the first laboratory experiments, some variations between the different colours were visible (Fig 9),

Table 5: Selected results at a water depth of 12 m and a camera distance of 1 m

	Laboratory	Underwater	Reconstruction
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			

but is not the case for the underwater condition as presented in Table 5.

The second general observation is that bright colours appear darker in the reconstruction than under laboratory conditions – best seen in the red colour patches nos. 12 and 15. Therefore, there is a tendency for low pixel values to be overestimated and high pixel values to be underestimated.

The high red content in the colour patches is reconstructed well. Patches nos. 1, 6, 10, 11, 12 and 15 contain a significant amount of red, and nos. 6 and no. 12 show especially good results for the red content. No. 11 shows some grey instead of orange, but this indicates problems in the combination of the channels and not in the red channel itself. In patches nos. 1 and 10, the red content is smaller and a degradation in the quality of the reconstruction can be observed immediately. This result is interesting considering that the average error in the red channel is the highest of all three colour channels.

The green content in the colour patches is also well reconstructed. Patches nos. 4, 7, 8, 9 and 14 contain a significant amount of green, and the bright green patch no. 9 is excellent. The strong green patches nos. 7 and 14 show problems: except for some red, the resulting colour patches appear also green in the reconstruction even though the reconstruction is darker. In patches nos. 4 and

no. 8 the green content is slightly lower and the resulting colour is not mediocre. However, the green content can clearly be seen in the reconstruction, so the problem does not lie with the green content. For patch no. 8, the mean colour in the reconstruction actually corresponds to the colour of the patch under white light. Green light has the longest penetration depth in coastal water, and the green colour channel shows the best overall results with a lower RMSE than the other two colour channels. Hence, the results are to be expected.

Regarding the blue content in the colour patches, high blue values are found in patches nos. 4, 5 and 8. Patch no. 5 is a dark blue in the laboratory with no red and green content and a small amount of blue. Patches nos. 4 and 8 are much brighter and therefore contain more blue. From these general observations, an underestimation of the blue value can be expected and is also apparent in Table 5. In patch no. 8, the blue content is higher and the reconstruction of the blue value is very good. One may observe some grey in patch no. 8, but since the green and blue values are both predicted well, the problem is more likely to be found in the red channel. A lower content of red is expected compared to the other two colours. As there is some grey, the red value is overestimated in this case.

Generally, the visual results show that the reconstruction of the colours are good even in the severe optical conditions under which the images are obtained. For the deeper water depths, the results are even better and it can be expected that for clear ocean water the errors are much smaller than the ones obtained in the experiment.

5. Conclusion and outlook

This study tested the applicability of different learning methods for underwater image reconstruction. It showed that the KNN and SVM methods are excellent choices to perform this task in the laboratory, subject to the image data being pre-processed appropriately. It also showed that reflections resulting from a focused light source reduced the performance of the learning machines. In addition, the distinction between red objects and black or dark grey objects is very difficult and sometimes leads to inappropriate colours in the reconstructed image.

During experiments in a diving basin, the KNN gave good results for short distances between the object and the camera, and for shallow water depths in the red channel. For longer distances and deeper water depths, and for the other colour channels, SVMs were the best choice for the reconstruction of

the colour under white light from the underwater images. Furthermore, the reconstruction under real conditions was much more difficult than in the laboratory due to the harsh conditions.

It was shown that even under extreme conditions the reconstruction of the colours was good, although the RMSEs of the colour channels were high. It could be seen that bright colours appear darker in the reconstruction and the high colour contents were easier to reconstruct.

The approach presented currently uses the original image without any additional filtering except for data smoothing. In the next stage, image pre-processing algorithms should be included to remove reflection effects and to smooth the brightness in the image. Additionally, the pre-processing after Bazeille et al. (2006) should be tested to increase the overall image quality. It may also be of interest to include information of neighbouring pixels for the calculation of the colour to remove single pixels that were not transformed correctly.

Furthermore, the setup should be improved by using a distributed light source. The laboratory results show that a focused light presents addition obstacles for the algorithms. In water with a significant number of suspended particles (like in the diving basin used for the experimental verification), focused light for illumination is also not possible as the light is directly reflected by the particles.

The procedure can also be used for other applications, where other ranges of wavelengths are used, to obtain the image as seen under white light from data under different lighting conditions. An example would be night sensing equipment like night-vision glasses or night vision devices in cars.

Finally, the experimental setup should be implemented into a real system, e.g. underwater robots and research submarines, and tested further under various conditions. Right before deployment, the learning machines should be trained for the onsite conditions. This should be done with a suitable known colour pattern as, for instance, the one used in this study in the underwater experiment.

Acknowledgment

The authors would like to thank Prof. Dr. rer. nat. Johannes Gottschling from the Chair of Mathematics for Engineers at the University of Duisburg-Essen for his insights into statistical learning. Furthermore, the authors thank the TauchRevier Gasometer for allowing the use of their diving facilities and Andreas Scholz and Stefan Westermaier for spending a Sunday in cold water taking pictures for the validation of the algorithms.

References

- Åhlen J. (2005). *Colour Correction of Underwater Images Using Spectral Data*. PhD. Dissertation, Uppsala Universitet.
- Aleksander I and Morton H. (1995). *An Introduction to Neural Computing*, second edition. Boston, MA: International Thomson Computer Press, 288 pp.
- Bayes T. (1793). An Essay towards solving a Problem in the Doctrine of Chances. *Phil. Trans.* **53**: 370–418.
- Bazeille S, Quidu I, Jaulin L and Malkasse J-P. (2006). Automatic Underwater Image Pre-Processing. In: Proceedings of Characterisation du Milieu Marin, 16–19 October, Brest.
- Ben-Gal I. (2007). Bayesian Networks. In: *Encyclopedia of Statistics in Quality & Reliability*. Ruggeri F, Kenett R and Faltin F, (eds.). Hoboken, NJ: Wiley & Sons.
- Carothers K. (2016). <http://oceanexplorer.noaa.gov/explorations/04deepscope/background/deeplight/media/diagram3.html>. <last accessed 17th May 2016>.
- Celebi AT and Ertürk S. (2012) Visual enhancement of underwater images using Empirical Mode Decomposition. *Expert Systems with Applications* **39**: 800–805.
- Chen Z, Wang H, Shen J, Li X and Xu L. (2014). Region-specialized underwater image restoration in inhomogeneous optical environments. *Optik, International Journal for Light and Electron Optics*. **125/9**: 2090–2098.
- Cooper GF and Herskovits E. (1992). A Bayesian method for the induction of probabilistic networks from data. *Machine Learning* **9**: 309–347.
- Cortes C and Vapnik V. (1995). Support-Vector Networks. *Machine Learning* **20**: 273–297.
- Drucker H, Burges CJC, Kaufman L, Smola A and Vapnik V. (1997). Support Vector Regression Machines. *Advances in Neural Information Processing Systems* **9**: 155–161.
- Ferreira C. (1996). Designing Neural Networks Using Gene Expression Programming. In: Abraham A, de Baets B, Koeppen M and Nickolay B. (eds). *Applied Soft Computing Technologies: The Challenge of Complexity*. Berlin and Heidelberg: Springer-Verlag. 517–536.
- Fix E and Hodges JL. (1951). *Discriminatory analysis, nonparametric discrimination: Consistency properties*. USAF School of Aviation Medicine Technical Report 4.
- Friedman N, Geiger D and Goldszmidt M. (1997). Bayesian network classifiers. *Machine Learning* **29**: 131–163.
- Gámez JA, Mateo JL and Puerta JM. (2011). Learning Bayesian networks by hill climbing: efficient methods based on progressive restriction of the neighborhood. *Data Mining and Knowledge Discovery* **22**: 106–148.
- Heckermann D, Geiger D and Chickering DM. (1995). Learning Bayesian networks: The combination of knowledge and statistical data. *Machine Learning* **20**: 197–243.
- Iqbal K, Salam RA, Osman A and Talib AZ. (2007). Underwater Image Enhancement Using an Integrated Colour Model. *International Journal of Computer Science* **34**: 529–534.
- Jaffe JS. (1990). Computer modeling and the design of optimal underwater imaging systems. *IEEE Journal of Oceanic Engineering* **15**: 101–111.
- Jaynes ET and Bretthorst GL. (2003). *Probability Theory: The Logic of Science: Principles and Elementary Applications*. Cambridge: Cambridge University Press, 753 pp.
- Lee D, Kim G, Kim D, Myung H and Choi H-T. (2012). Vision-based object detection and tracking for autonomous navigation of underwater robots. *Ocean Engineering* **48**: 59–68.
- Lewicki P and Hill T. (2005). *Statistics: Methods and Applications*. Tulsa: StatSoft, Inc.

- MathWorks Inc. (2016). MATLAB <http://de.mathworks.com/help/vision/ref/matchfeatures.html>. <last accessed on 17th May 2016>.
- McCulloch W and Pitts W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* **5**: 115–133.
- Mobley CD. (2004). *Light and Water: Radiative Transfer in Natural Waters*. Office of Naval Research.
- Rojas R. (1996). *Neural Networks: A Systematic Introduction*. Berlin and Heidelberg: Springer-Verlag. 502 pp.
- Smith RC and Baker KS. (1981), Optical properties of the clearest natural water (200–800 nm). *Applied Optics* **20**: 177–184.
-