THE UNIVERSITY OF CHICAGO

CUTTLEFISH CAMOUFLAGE QUANTIFICATION VIA NOVEL NEURAL NETWORK APPROACHES AND HYPERSPECTRAL IMAGING

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
AND
THE FACULTY OF THE DIVISION OF BIOLOGICAL SCIENCES
AND THE PRITZKER SCHOOL OF MEDICINE
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN BIOPHYSICAL SCIENCES

BY
AMAR RISBUD

CHICAGO, ILLINOIS
DECEMBER 2022

To my parents, Dilip and Anuradha Risbud, who showed me that learning is a compass.

"The old god still remembered all his tricks,
and first became a lion with a mane,
then snake, then leopard, then a mighty boar,
then flowing water, then a leafy tree.
But we kept holding on: our hearts stood firm."

- Menelaus on the capture of Proteus, *The Odyssey*
Homer transl. Emily Wilson

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would first and foremost like to thank my thesis committee, Dr. Stephanie Palmer, Dr. Mark Westneat, Dr. Roger Hanlon, Dr. Steven Shevell, and Dr. Wei Wei not only for their academic wisdom, but also for their empathy and support when the path forward was not as sure. I could not have asked for a better dual mentorship than Dr. Palmer and Dr. Westneat, Dr. Hanlon taught me everything I know about cephalopod ecology and provided the amazing image set used in this thesis, Dr.was a great resource with respect to color vision, and Dr.offered a wealth of guidance with respect to human vision. I would like to additionally thank Dr. Kendra Buresch, an invaluable member of the Hanlon group. I would also like to thank Dr. Leslie Osborne. Dr. Osborne and her student, Neerav Goswami, provided the human eye-tracking dataset we used to validate the metric proposed in this thesis. Dr. Osborne's knowledge of human psychophysics was also an amazing resource when brainstorming experiments or directions of analysis. I also owe a tremendous debt of gratitude to the post-docs in the Palmer lab, chief among them Dr. Siwei Wang and Dr. Eyal Nitzany. Dr. Wang's neural network prowess was indispensable, and Dr. Nitzany introduced me to the idea of texture synthesis when I first started in Dr. Palmer's lab. I would also like to thank the entire Palmer group for their social and academic support over the years. Finally, I would like to thank the entire staff of the Biophysical Sciences Program, particularly Michele Wittels, Juliana Feder, Candice Lewis, Adam Hammond, Tobin Sosnick, and Greg Engel, for offering an essential administrative support structure throughout my graduate career.

Of course I would like to thank my family: my parents, Dilip & Anuradha Risbud, and my sister, Rashmi Risbud, for the motivation and support they have consistently provided me throughout my life. I would also like to thank my fellow UChicago students Taryn Serman, Evan Kiefl, Devin Harrison, Vaughn Spurrier, & Steven Redford for being amazing friends and colleagues. I would also like to thank my roommate and friend Dr. Scott Trinkle for our mutual love of '80s movies and Talking Heads. I would also like to thank the entire Chicago

comedy community and the staff at Cafe Mustache for making sure I had a life outside of graduate school. I would like to thank my best friend Keenan Lloyd, for always being there for me, telling me hard truths when I needed to hear them, and laughing at Teachers' Lounge episodes with me. Finally, I would like to thank my little cat Paulie, for being a constant (yet clueless) companion over these past three years.

ABSTRACT

Rapid adaptive camouflage is a critical defense mechanism for cephalopods. The characterization of cephalopod camouflage has thus far been reserved almost exclusively to qualitative descriptions, and research on camouflage quantification remains in a nascent state. Qualitative characterizations do not capture the full multifactorial nature of camouflage, nor do they provide a comprehensive metric by which the degree and effectiveness of cephalopod camouflage from the perspective of a given predator can be quantitatively measured in a given scene. Here, I propose a "texture distance" metric that integrates lower and higher dimensional visual features to give a pixel-wise read out of the similarity between a cephalopod's texture and its background texture. This metric is based on a previously developed algorithm that utilizes an artificial neural network to perform texture synthesis. Such a quantifying method would allow researchers to gain more insight into the amount of evolutionary pressure camouflage might exert on predator visual systems. The proposed metric is developed, validated, and used alongside other analyses to investigate the search strategies of human subjects in a camouflage detection task, especially with respect to the difference between human subjects who were successful or unsuccessful in the search task. Furthermore, hyperspectral images (HSI) of cephalopods under natural lighting conditions in the wild were used to measure chromatic and luminance discriminability in the visual color space of a subset of real natural fish predators. This HSI analysis suggests sophisticated color-matching across predator types, while also suggesting cephalopods are more detectable via changes in luminance. This result is compatible with the results of previous studies, but opens the door to exciting new research possibilities.

CHAPTER 1

INTRODUCTION

It seeks its prey by so changing its colour as to render it like the colour of the stones adjacent to it; it does so also when alarmed. By some the sepia is said to perform the same trick; that is, they say it can change its colour so as to make it resemble the colour of its habitat.

-Aristotle, *History of Animals* transl. D'Arcy Wentworth Thompson

Since the earliest characterization of cephalopods by Aristotle in his *History of Animals*, their camouflage capabilities have consistently captured the attention of researchers. Moreover, the scientific investigation of this camouflage has also consistently challenged researchers. The rapid, adaptive nature of camouflage, the diversity of body patterns, and the startling intelligence of the animals themselves raise a plethora of difficult questions. How is this camouflage controlled, what structures help produce it, and what makes an animal choose a specific body pattern at a specific time, and in a specific environmental context? While these questions of cephalopod neurobiology and the nature of camouflage-producing organs are certainly important and have been pursued by various researchers, they focus primarily on the *production* of camouflage while the question of *predator perception* has been left relatively under-studied: how can one begin to describe the camouflage patterns themselves, and more critically, can one do so from the perspective of natural predators?

This thesis aims to contribute to this burgeoning research on camouflage quantification. The central goals of this dissertation are (1) to propose a method for quantifying cuttlefish detectability in natural scenes, (2) to validate this method and gain insight into predator search strategies using eye-tracking data collected from human subjects searching for camouflaged cuttlefish, and (3) to investigate detectability of camouflaged cuttlefish in the visual spaces of real natural predators using hyperspectral imaging. In Chapter 2, we propose

the new "texture distance" metric for quantification based on neural network-driven texture synthesis [1]. This metric integrates lower and higher order features to quantify the textural similarity between an animal and its background. In Chapter 3, I validate the metric using a rich human eye-tracking dataset, and demonstrate the metric's utility in not only predicting the search strategies of visual subjects, but also identifying differences between successful and unsuccessful searches. In Chapter 4, I further examine the use of hyperspectral images (HSI) to simulate predator visual perspectives, building on foundational work by Chiao, et al. and Dr. Martin Stevens [2, 3] by using HSI images taken in the wild under natural conditions, using real predator spectral sensitivities.

The development of this metric, along with the HSI analysis, can then aid in the investigation of how the vision of predators that share ecological niches with cephalopods has evolved to contend with camouflage, a path of inquiry that surpasses the utility of the aforementioned qualitative, subjective categorizations of camouflage type. Additionally, such a method could be applied in any biological context in which camouflage plays a role, whether marine or otherwise. In a broader sense, it is easy to imagine that the ability to quantify the effectiveness of a foreign material blending onto a substrate would be relevant in the context of medical imaging, materials science, and other fields that deal with the visual continuum of conspicuity to camouflage.

CHAPTER 2

TEXTURE DISTANCE METRIC FOR CAMOUFLAGE QUANTIFICATION VIA TEXTURE SYNTHESIS

This work was completed in collaboration with Dr. Siwei Wang, who implemented the output of the VGG-19 loss function in Python, and Dr. Stephanie Palmer.

2.1 Abstract

Here, I propose the use of a popular texture synthesis model as the basis for a metric for cephalopod camouflage quantification. Given an image of cephalopod camouflage, calculating the loss function of this model between two inputs, (1) a reference patch on the cuttlefish, and (2) every other equally sized patch in the image yields a "texture distance" heatmap of the scene. This heatmap gives the textural similarity between every pixel in the scene and the camouflaged cephalopod. While research in the field of camouflage quantification is burgeoning and alternative quantification metrics do exist, the advantages of this texture distance metric are its compatibility with qualitative camouflage standards, its integration of lower and higher order features, and its potential applicability to a range of camouflage contexts and orientations.

2.2 Introduction

Cephalopods such as the common cuttlefish (*Sepia officinalis*) rely on rapid adaptive camouflage for primary defense [4, 5, 6]. In contrast to the static camouflage of many animals, cephalopods are capable of actively and rapidly changing their overall camouflaged body pattern in a wide range of natural habitats. These body patterns include dynamic color, contrast, luminance, pattern, and 3D skin texture. The complexity and diversity of

cuttlefish camouflage patterns and contexts is astonishing, and extremely visually impressive. Cuttlefish are prey to many predators (representing a wide array of visual capabilities) throughout many varied ecological niches, being represented in many stomach-content studies [7, 8, 9, 10, 11, 12, 13] and having at least one predator (*Serranus cabrilla*) that has been observed repeatedly targeting them in the wild [4]. Therefore, given the complexity of cuttlefish camouflage, predators must invest substantial resources in order to break camouflage, i.e. positively identify their prey when it is blending into the environment. Measuring the visual properties of regions where visual predators look when they search for a camouflaged cuttlefish in an image can better elucidate the nature of camouflage, specifically with respect to where it works best and how it can fail. Understanding the precise effectiveness of camouflage in a given scene, particularly in the visual space of a given predator, can provide insight into the state of the evolutionary arms race between predator and prey in a given ecological niche.

As such, camouflage has been of great interest to researchers, although a more complete understanding of its function and generation has only recently begun to be achieved. Before continuing, it is important to introduce the history and current state of research in camouflage, with particular attention to how our proposed metric contributes to its advancement.

Among the earliest sophisticated descriptions of camouflage in modernity was the study by the artist Abbott Handerson Thayer, *Concealing-coloration in the animal kingdom* [14]. Thayer's development of the principle of countershading, whereby "Animals are painted by Nature darkest on those parts which tend to be most lighted by the sky's light, and vice versa" [14], had a large influence on the landmark study by Hugh B. Cott, *Adaptive coloration in animals* [15]. Cott incorporated and expanded upon Thayer's characterization by introducing four basic requirements for successful camouflage, namely 1. color resemblance, 2. obliterative shading, 3. disruptive coloration, and 4. shadow elimination [15]. Thayer and Cott [14, 15] established general principles for camouflage across the animal kingdom, but in more recent

Figure 2.1: Cephalopod camouflage is extremely visually compelling, and can be deployed in an astonishingly wide range of environmental contexts. The images above show the common cuttlefish (*Sepia officinalis*) in various states of camouflage. Images from Turkey by Dr. Roger T. Hanlon and team.

years, numerous advances in understanding cephalopod camouflage in particular have been made.

The characterization of cephalopod defense behavior and the UMD (uniform-mottle-disruptive) schema for categorizing cephalopod body patterns developed by Hanlon & Messenger [4, 5] was a major organizational step forward in the field. Since this study concerns the quantification of cuttlefish camouflage, an overview of the behaviors and body patterns described by Hanlon & Messenger [4, 5] would be instructive. Though specific strategies vary widely across species, cephalopod defense can be broadly separated into two classes: primary defense and secondary defense [5]. Primary defense mechanisms are employed in order to avoid detection or recognition by predators, while secondary defense mechanisms are primarily evasive or confusing behaviors aimed at frustrating pursuit by a predator following recognition. Here, detection is defined as the reflected light from the cephalopod registering in the predator visual system as a stimulus that is distinct from its background, with the condition that the predator does not recognize this visual stimulus as corresponding to the presence of a specific prey animal. Meanwhile, recognition is defined as the predator's correct association of that distinct visual stimulus to a prey animal following detection. This thesis deals with primary defense strategies, so a detailed description of secondary defense is not provided here, several papers on secondary defense in cephalopods have been published, and some aspects of secondary defense have been studied in a laboratory setting [16].

Primary defense strategies, otherwise referred to by the umbrella term "camouflage", can in turn be subdivided into three major categories: crypsis, masquerade, and mimicry [5]. Crypsis is perhaps the most important subcategory in the context of this thesis project, since it includes those behaviors which allow the cuttlefish to avoid detection by replicating certain features of its natural surroundings to minimize its distinguishability from those surroundings. Crypsis itself is composed of a number of sophisticated substrategies, including background matching [5], disruptive patterning [5, 3, 17], distractive markings [5, 18], coun-

tershading/counterillumination [5, 19, 20, 21, 22], shadow elimination [4, 5], transparency [5, 3], night camouflage [5, 23], and motion camouflage [5, 24]. Of course, these strategies can intersect and overlap in meaningful ways, so it is helpful to understand them as complementary rather than as entirely distinct. It is important to note that the sophistication of primary defense strategies arises in large part from the production of body patterns through the coordinated neural regulation of chromatophores, leucophores, iridophores, and muscles in the skin of cephalopods [5].

Perhaps the most important substrategy of crypsis in the context of this thesis project is background matching, since the indiscriminability of a cuttlefish from its natural background is the goal of this strategy, and *measuring* that discriminability is the goal of this project. Background matching describes the strategy by which a cephalopod minimizes its distinguishability from a natural background by matching the color, brightness, contrast, roughness, and pattern of its skin to that of the background [5]. Based on their studies of cuttlefish camouflage, Hanlon & Messenger [4] have proposed three major types of patterns cuttlefish typically employ when stationary on a natural background. The first pattern type, Uniform, is low-contrast and homogeneous, intended to facilitate camouflage on uniformly colored, relatively smooth textures, like beds of sand. An example of a uniform pattern is shown in the top right panel of Figure 2.1. The second pattern type, Mottle, is chromatically heterogeneous, and consists of average-contrast, light and dark, roughly ovular "components". Each of these components can be thought of as an ordered cluster of coloration organs and structures like those discussed above. "Light" describes components in which chromatophores are retracted, thereby exposing the comparatively light skin or other light elements underneath. Conversely, "dark" describes components in which chromatophores are expanded, occluding these lighter elements with colors corresponding to the chromatophore pigments. The Mottle pattern is employed to facilitate camouflage on chromatically heterogeneous textures whose individual units range in spatial scale from small to

medium. An example of the Mottle pattern is shown in the bottom left panel of Figure 2.1. Research has suggested that mottle body patterns are associated with backgrounds characterized by low spatial frequencies. Unevenly colored beds of roughly equally small pebbles or rocks might require a mottle pattern. Finally, the third pattern type, Disruptive, is also chromatically heterogeneous, and consists of high-contrast, light and dark components. The range of spatial scale, as well as the variety of shapes, of the components that make up this pattern is greater than that of the mottle and uniform patterns. Past studies have suggested that disruptive body patterns are associated with backgrounds characterized by both low and high spatial frequencies [25, 26]. An example of a Disruptive body pattern is shown in the top left and bottom right panels of Figure 2.1. In addition to employing these three pattern types for background matching, cuttlefish are also able to modify the smoothness or roughness of their skin to match that of the background through appropriate extension of skin papillae [5]. Additionally, cuttlefish can improve the quality of the match to the background by arranging limbs in context-specific configurations [5].

It is worth noting that the utility of disruptive patterning extends beyond the purpose of matching to highly non uniform background textures [5]. Disruptive body patterns obscure the true outline of a cephalopod, render the outline unrecognizable, or create the impression of a false outline. These patterns are produced through a combination of high contrast shapes, many of which run perpendicular to the major axis of the cephalopod and terminate at its boundaries. While these patterns can be used in the context of matching to non-uniform backgrounds as discussed earlier, they can also be deployed in non-matching contexts to create visual stimuli that are distinct from the background but which obscure the outline of the cephalopod in such a way that its identity is not betrayed. In this way, detection occurs, but recognition is avoided. This is a critical point - the quantification metric introduced in this thesis certainly applies to uniform and mottle body patterns, but only applies to disruptive coloration to the extent that it is used as a background matching strategy that

retards detection. It is also important to address that the metric proposed in this thesis does not apply to the other two primary defense substrategies mentioned above: masquerade and mimicry [5]. Masquerade refers to a cephalopod's ability to imitate non-animal objects in the environment (imitation of rocks and coral in "moving rock" behavior of *O. cyanea* [24]), while mimicry refers to a cephalopod's ability to imitate other animals in the environment (imitation of flatfishes by *Thaumoctopus mimicus* [5]). These strategies do not involve the avoidance of detection, but rather the avoidance of recognition via the cephalopod's adoption of a false identity. The goal of the quantification metric proposed in this thesis is to measure how well a cephalopod avoids detection, so it should not be employed with respect to post-detection body patterns or strategies that only seek to retard recognition.

In addition to the qualitative UMD characterization described above, recent advances in cephalopod camouflage research have also attempted to approach camouflage *quantitatively*: is the development of a quantitative measure that reads out how effective camouflage is in a given scene possible? Past studies [25, 2] have tried to answer the question of quantification by examining the spatial frequency of the cephalopod's body pattern, or by measuring the difference in spectral frequency between a cephalopod's body pattern and its background. The spatial frequency analysis conducted by Barbosa, et al. [25] examined the effects of background contrast and spatial scale on cuttlefish body patterns, and resulted in an automated granularity program. In this program, passing an image of a cuttlefish through multiple band-pass filters of different spatial scales reads out a Fourier-transform style plot of spatial frequency vs. energy. The spectral analysis conducted by Chiao, et al. [2] used hyperspectral imaging to investigate the color match between background and animal, with the additional demonstration of how this color match analysis could be performed through the visual perspective of natural predators. The limitations of the spatial frequency and color-matching studies lie in the fact that each of these measurements relies on only a single factor. While the shape of the curve generated by the granularity program could indicate

the body pattern type, it does not give information regarding color or orientation. Similarly, while the spectral analysis can be used to determine whether the color of a cuttlefish matches the background, it disregards contrast information. Even if one were to integrate these two measurements, it is clear that considering only these factors is not sufficient to capture the complexity of cephalopod camouflage. Camouflage relies not only on matching spectral and spatial frequency, but also on replicating the directionality, complex edges, contours, and curves of the surrounding environment. More recently, a survey by Troscianko, et al. [27] has comprehensively compared a range of existing popular quantification metrics, in addition to proposing a new metric based on biologically relevant Gabor filters used for edge detection. In light of the restricted focus of the spectral and spatial frequency analyses, the survey by Troscianko, et al. [27] showcases the promise of more advanced quantification methods such as the Fourier-style SIFT analysis, the neural-net inspired HMAX method, and their proposed GabRat filter based on biologically relevant Gabor filters.

The metric proposed in this thesis, which will be referred to as the "texture distance metric", builds on these more complex, biologically inspired quantification metrics by fully embracing the tremendous power of deep neural networks. In recent years, this power has been demonstrated in the form of image classification [28], saliency predictors [29, 30, 31], and texture synthesis algorithms [1, 32]. It is upon the latter that this new method of measurement is built. Framing camouflage as a cuttlefish attempting to match the texture of its background, I propose a new metric for images of camouflage that gives a score at each pixel of textural similarity to the camouflaged animal. This metric is composed of a summation of feature losses from the output of a texture synthesis algorithm in which the two inputs are a reference patch from the camouflaged cuttlefish, and an equally sized patch from the background. The goal of texture synthesis algorithms is to generate a synthetic texture that matches a real input texture from a starting white noise image, so the feature losses are simply a measure of how texturally different the two input patches are. The texture synthesis

GabRat	"Average ratio of 'false edges' (edges at right angles to the prey outline) to 'salient edges' (edges parallel with the prey outline)." [27]
HMAX	"Breaks down a bank of Gabor Filter outputs into layers that describe patterns with some invariance to scale and orientation" [27, 34]
ImageNET	"An image database organized according to the WordNet hierarchy, in which each node of the hierarchy is depicted by hundreds and thousands of images. Frequently used to train image classification networks such as VGG-19." [33]
ReLU	A piecewise activation function enabling the learning of complex, nonlinear operations in deep neural networks.
Texture Synthesis	Algorithms that aim to match the texture of an output image to an input target texture.
VGG-19	A very deep convolutional neural network with 19 weight layers and 3x3 convolutional filters. [28]

Table 2.1: Glossary of key machine-learning terms.

algorithm utilized for the proposed metric is detailed in the following section. The use of a VGG-19 image classification model [28] pre-trained on over a million images of ImageNet [33] gives our metric a distinct advantage over past attempts to quantify camouflage. This sophisticated learning results in a wider range of filter types and complexity, building on the cortical structure of the Gabor-only HMAX method [27]. The proposed GabRat metric in Troscianko, et al. [27] focuses on the quantification of disruptive coloration. However, the texture distance metric may be amenable to quantifying background matching and disruptive coloration. Finally, while the Troscianko, et al. survey [27] simulates camouflaged prey on backgrounds, the validation of the texture distance metric is performed on images of real cephalopods camouflaging in real underwater scenes, under natural lighting conditions.

2.3 Methods

The metric proposed above relies on the texture synthesis algorithm developed by Gatys, et al. [1] This algorithm demonstrated how, with the use of the VGG-19 artificial neural network, a white noise image could be transformed into a synthetic sample of a real input texture. However, the idea of texture synthesis, the recent history of its development, and a precise definition of "texture" must be established before moving forward. When "texture" is mentioned in this thesis, it refers to the definition set forth in the seminal texture synthesis paper by Simoncelli & Portilla [35]: "Loosely speaking, texture images are spatially homogeneous and consist of repeated elements, often subject to some randomization in their location, size, color, orientation, etc." This definition of texture can be connected to the repeating components of the body patterns in the UMD characterization of cuttlefish camouflage. With the exception of certain conspicuous markings in disruptive coloration, the body patterns deployed by cuttlefish during background matching are spatially isomorphic, consist of repeating chromatophore components, and subject to variation based on environmental context and direct neural control over the intensity of patterning. So, I propose that a sophisticated texture synthesis algorithm could be used to measure the textural difference between a cephalopod's body pattern and the background, and thereby quantify camouflage.

The impetus underlying texture synthesis is the faithful synthetic reproduction of a real input texture using a set of statistics that approximates the structure of the early stages of visual perception [36, 35, 1]. Such a faithful reproduction would validate that statistical model as being representative of a visual system. This research can be traced from the early work by Julesz [36, 35], categorizing textures via the N-th order joint histogram of the image pixels, to the multiple spatial scale/orientation structure of Simoncelli & Portilla's [35] wavelet representation, inspired by kernels that model early mammalian visual processing. While the texture synthesis model developed by Simoncelli & Portilla [35] is a landmark, the algorithm developed by Gatys et al. [1] sought to address the limitations of its predecessor

by using the complex filters and non-linearities of a convolutional neural network (VGG-19) trained on ImageNet rather than a curated sample of linear filters [1, 28, 33]. According to Gatys, et al. [1], a convolutional neural network stands in for the whole ventral pathway of visual processing, moving beyond the early visual modeling of Simoncelli & Portilla [35].

Indeed, the sophistication of the VGG-19 network is demonstrated by its first-place performance in the classification and localisation portions of the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2014 [28]. Trained on more than 1 million images from the ImageNet database, VGG-19 is an artificial deep convolutional neural network - consisting of layers of convolutional filters interspersed with rectifications and pooling operations to simulate human visual processing [28]. The Gatys, et al. [1] texture synthesis model seeks to minimize the mean squared error (MSE) between the Gram matrices of the activation maps of selected layers of the VGG-19 network for the input and synthetic texture. This MSE is given by

$$\mathcal{L}_{style}(\alpha, \chi) = \sum_{\ell=0}^{L} w_l E_l, \tag{2.1}$$

where w_l are the weights of each layer when calculating the MSE, and α and χ are the style image and the generated output image. E_l is the contribution of a single layer to the MSE and is given by

$$E_l = \frac{1}{4{N_l}^2{M_l}^2} \sum_{i,j} (G_{ij}^l - A_{ij}^l)^2.$$

N_l is the number of distinct filters in the layer, and M_l is the area of the feature maps of those filters. G_{ij}^l and A_{ij}^l are the dot products of feature maps i and j in layer l for the input image and the generated image, respectively.

It is precisely this MSE statistic that forms the basic unit of our texture distance metric, as will be explained in the following section. Gatys, et al. [1] found that constraining the MSE by all layers up to the fourth pooling layer resulted in high-quality textures, although the

number of parameters was extremely high. As a compromise, in order to consider multiple layers across spatial scales while keeping the number of parameters relatively low, the MSE is calculated using the convolutional layers `conv1_1`, `conv2_1`, `conv3_1`, `conv4_1`, and `conv5_1`. The validity of this set of layers was also demonstrated by its use in the later neural style transfer paper by Gatys, et al. [32], in which these layers were used to transfer artistic style with great sophistication.

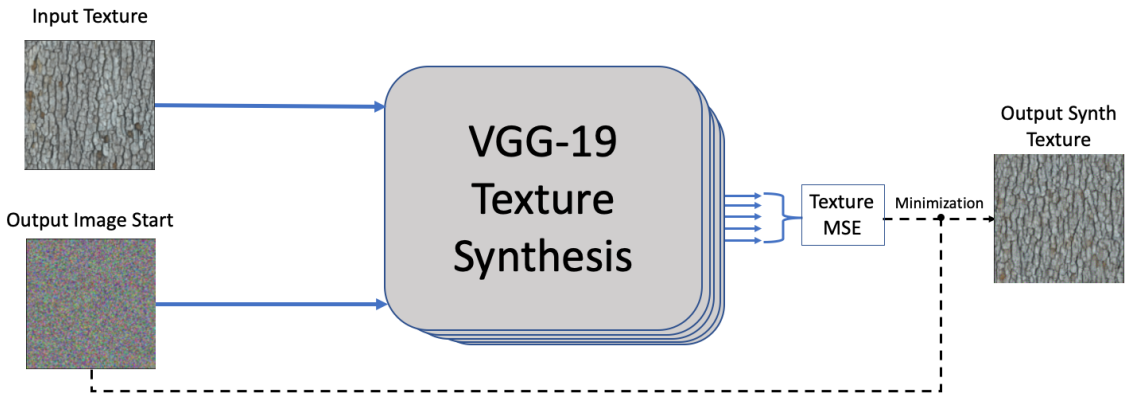


Figure 2.2: Texture synthesis uses multiple convolutional layers in the VGG-19 image classification deep neural network to minimize a loss function to yield an output synthetic texture that matches the input.

The general structure of the texture synthesis model is shown in Figure 2.2, along with example input and generated synthetic textures. The output start image, in this case random white noise, is changed via gradient descent until the MSE between the output and the input texture is minimized [1].

The starting point for the proposed metric is a JPEG image of a cephalopod resting on a natural or artificial substrate. As an example, an image of a cuttlefish camouflaging in a coral outcropping will be used, as shown in Figure 2.4. The first step in calculating the proposed metric is to select a representative image from within the boundaries of the cephalopod. Next, the MSE between this representative cephalopod patch and an equally sized patch in the image is calculated using a Python implementation of the VGG-19 neural network. As explained above, the MSE is calculated using the convolutional layers `conv1_1`,

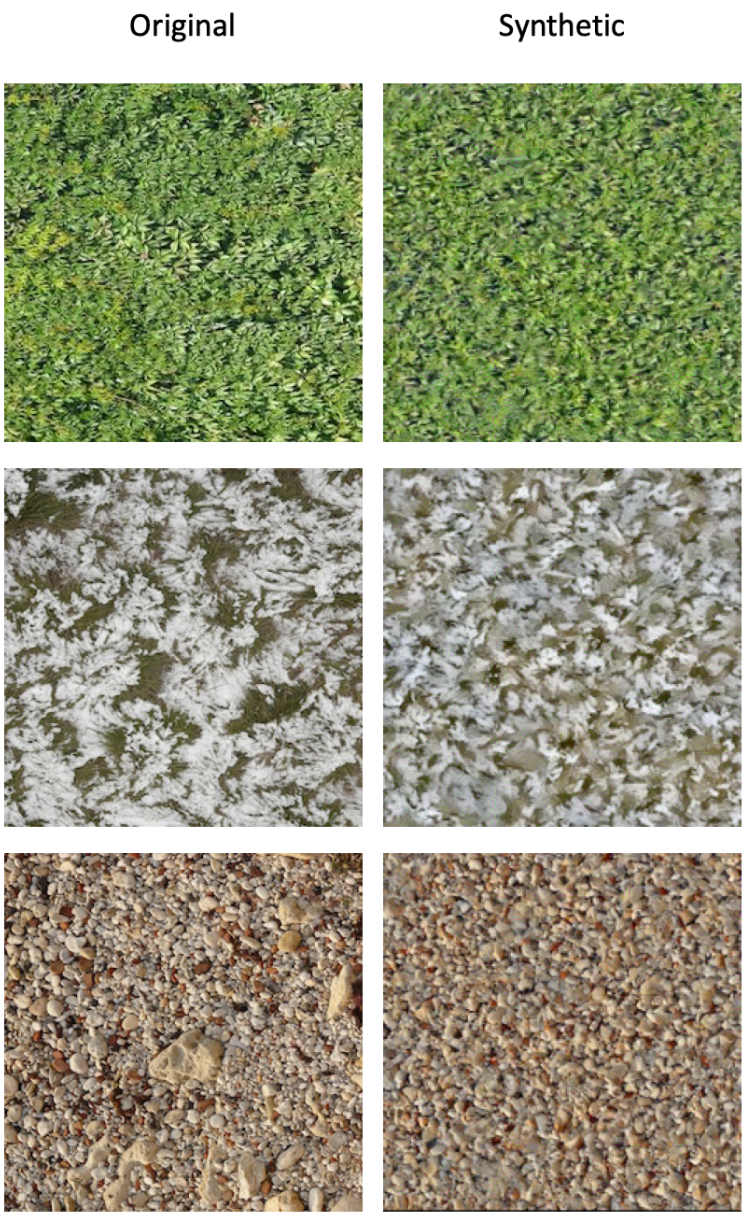


Figure 2.3: Texture synthesis uses multiple convolutional layers in the VGG-19 image classification deep neural network to minimize a loss function to yield an output synthetic texture that matches the input.

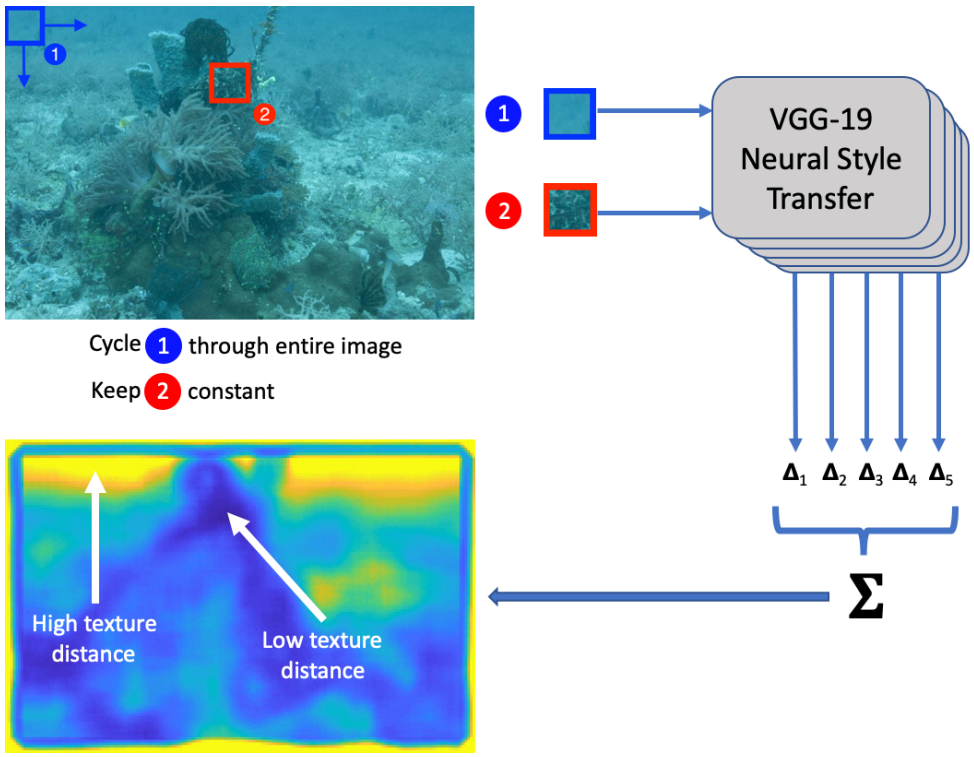


Figure 2.4: Texture distance heatmaps are generated from images by calculating the texture distance between a reference patch on the cephalopod and background patches taken at five-pixel steps. In the resulting heatmap, each pixel gives the texture distance from the cephalopod to that region of the image. Image from Turkey by Dr. Roger T. Hanlon and team.

`conv2_1`, `conv3_1`, `conv4_1`, and `conv5_1`, as in Gatys, et al. [1] The patch size was chosen to be 69x69 pixels in size, which satisfies the requirements of the convolutional filters selected, and on average adequately samples the texture of the camouflaged cuttlefish in a given image. By keeping the representative cephalopod patch constant, changing the equally-sized background patch by a fixed step size (five pixels), and cycling in this way through the entire image, a heatmap of texture distance values that shows the stylistic dissimilarity between any given point in the image and the representative cephalopod patch can be built. The image set used consist of 105 images, and were collected by Dr. Roger Hanlon and his team in Turkey, and depicts the common cuttlefish *Sepia officinalis* camouflaging in various reef scenes.

2.4 Results

A number of texture distance heatmap examples are also given in Figure 2.5. The dark blue regions represent areas that have a small texture distance from the cuttlefish, i.e. regions that are stylistically similar to the representative patch from the cuttlefish. Conversely, the bright yellow regions represent areas that have a large texture distance from the cuttlefish, i.e. regions that are stylistically dissimilar to the representative patch from the cuttlefish. In the case of theoretically perfect camouflage on a totally uniform background, the heatmap would be completely blue, since a representative sample from a perfectly camouflaged animal would have the same local feature correlations and complex features as a random sample from the background, and would therefore yield a style loss of zero. I therefore propose the following hypothesis: that any agent searching for a cephalopod will search the regions where the texture distance between cephalopod and background is low.

2.5 Discussion

In this chapter, I propose the use of a popular texture synthesis algorithm as the basis for a quantification metric for cuttlefish camouflage. I demonstrate how, given an input image of an underwater scene containing a camouflaged cephalopod, the loss function of this texture synthesis algorithm can be used to generate a heatmap that shows the textural similarity of any patch in the image to an equally sized reference patch sampled from the cephalopod. The "texture distance" metric developed here aims to address the restrictions of previously developed quantification metrics. Initial forays into camouflage quantification focused on the measurement of a single visual factor (color or spatial frequency). While a metric such as the granularity analysis is able to categorize pattern type based on the shape of a spatial frequency curve, it is restricted in that it requires the removal of color information and is sensitive to orientation, i.e. the cephalopod in an image needs to be warped and oriented

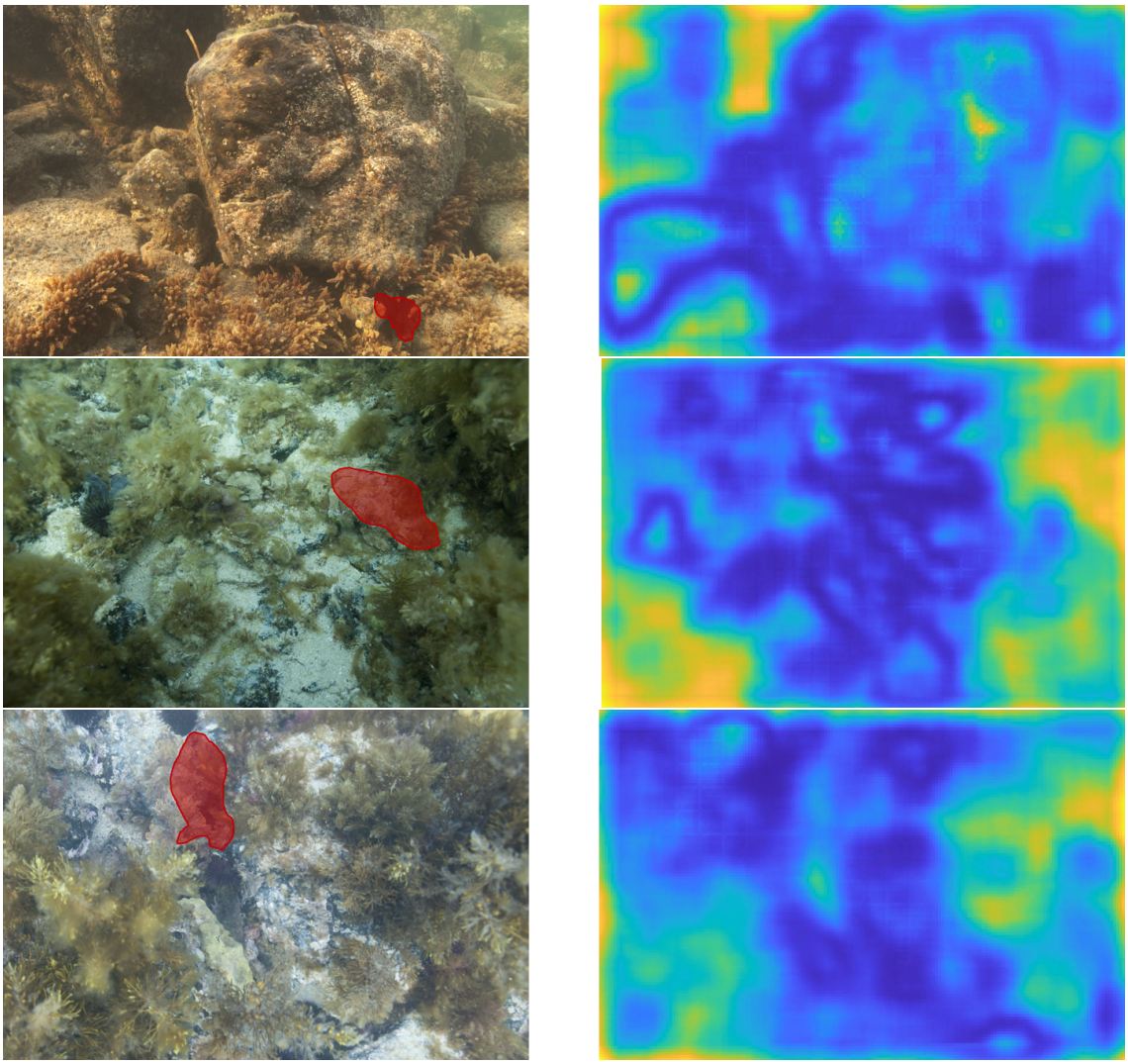


Figure 2.5: Texture distance heatmaps (between cuttlefish and background) are shown alongside their corresponding images for a number of examples, with cuttlefish locations masked in red. Note the diversity of possible camouflage patterns and environments, and the intuitive textural similarities, differences, and features highlighted in the heatmaps. Heatmaps are generated for the entire image dataset discussed in the following chapter (see Chapter 3). Image from Turkey by Dr. Roger T. Hanlon and team.

appropriately in order to be input to the granularity program. Furthermore, the granularity analysis takes into account a curated selection of spatial frequency bands. The texture distance metric, by contrast, integrates lower and higher order visual features due to the depth of the neural network. The analysis of spatial frequency does not have to take place at the expense of color, and vice versa. In fact, the texture distance metric not only takes into account these two key features, but rather a broad range of diverse visual features as a result of the vast number of filters in the neural network. It is this same characteristic that allows the texture distance metric to be invariant to orientation - adjustment of the direction or shape of the body pattern in the image is not necessary prior to input. Furthermore, the range of spatial scales represented by the multiple convolutional layers means that, as long as a reference patch from the cuttlefish is large enough to meet Simoncelli & Portilla's [35] definition of texture, the texture distance metric is not sensitive to the size of the cephalopod in the image. In this way, the texture distance metric has some advantage over the single-factorial measures developed in the initial stages of camouflage quantification research.

It must be noted, however, that recent research into camouflage quantification has progressed beyond single-factorial measures, so it is important to understand where the texture distance metric stands in relation to these newer, more complex and biologically inspired metrics. The restrictions of the GabRat and HMAX method covered in the survey by Troscianko, et al. [27] were briefly discussed in the Introduction, but it is important to revisit them here. While the HMAX method addresses the issue of sensitivity to orientation and scale, and is inspired by neural architecture, it is restricted to the use of Gabor filters. By contrast, the VGG-19 neural network has an abundance of diverse filters in each layer that are learned through training on ImageNet. This training means that while some of the filters inside the proverbial "black box" of the network may be similar to Gabor filters, others may be much more complex and targeted to even higher-order shapes and contours that have not yet been characterized. It is worth repeating the point that the neural network's depth is

important, since it greatly increases the descriptive resolution of the texture distance metric. The GabRat metric is similarly based only on Gabor filters, and is applied only to cases of disruptive coloration, whereas the texture distance metric could be applicable to any instance of background matching. So, in addition to moving beyond the single-factorial measures of camouflage quantification, the texture distance metric also addresses some of the restrictions of more recent biologically-inspired, complex quantification metrics.

Where might such a quantification metric prove useful? It is important to note that the texture distance metric is by no means restricted to the study of cuttlefish camouflage, but is amenable to the study of any type of camouflage by any type of organism. This equips the texture distance metric for the quantification of camouflage by other underwater organisms, lizards, and insects. Beyond the scope of evolutionary biology, the texture distance metric can also be applied to the study of adaptable materials and cloaking technologies, namely for the evaluation of their ability to blend into a background substrate. The application of convolutional neural networks to tumor identification has also been studied [37], and it is easy to imagine how an optimized or modified version of texture distance heatmaps generated earlier might be useful in the context of medical imaging. In sum, the texture distance metric is potentially applicable to any field in which the spectrum of conspicuity to camouflage is important.

CHAPTER 3

EXAMINING THE RELATIONSHIP BETWEEN THE TEXTURE DISTANCE METRIC AND HUMAN GAZE DATA

This work was completed in collaboration with Dr. Leslie Osborne and Neerav Goswami, who developed the methodology and execution of the eye-tracking task, and Dr. Siwei Wang and Dr. Stephanie Palmer, who assisted with data analyses.

3.1 Abstract

Here, I validate a texture distance metric, which was developed from a popular texture synthesis model for the purpose of quantifying cephalopod camouflage. This validation uses a rich eye-tracking dataset, collected on a group of human subjects performing a search task to locate and positively identify a camouflaged cephalopod in a natural underwater scene. From this eye-tracking data, I conclude that any subject searching for a cephalopod will spend the majority of their time fixating on regions of low texture distance. Furthermore, I conclude that subjects who succeed in identifying the cephalopod spend a lower proportion of time in low texture distance regions than those who failed to identify the cephalopod. Since these results suggest a reliable relationship between the texture distance metric and predator behavior (assuming humans as a model for a general cuttlefish predator), it is reasonable to conclude that the proposed texture distance metric is not only a valid quantification method, but also shows promising potential as a predictor of human gaze.

3.2 Introduction

Visual fixation data from human subjects is a rich way to investigate the relationship between our proposed texture distance metric (see Chapter 2) and the search strategies of natural predators. The purpose of such an investigation is twofold. First, if a meaningful

relationship exists, the proposed metric is validated as a promising camouflage quantification measure, since one could in theory infer the relative difficulty of a search task from the statistics of texture distance values in a given scene. Second, such an investigation would elucidate the strategies natural predators employ when searching for cephalopods, and particularly the differences between strategies that succeed and strategies that fail. Furthermore, humans are a good stand-in for natural predators, as humans are not only extremely skilled at pattern recognition and visual search, but easily accessible and instructable. By measuring the visual properties of regions where human subjects look when they search for a camouflaged cuttlefish in a curated set of images, perhaps the nature of camouflage can be made more clear, specifically with respect to where it works best and how it can fail. In the absence of a difficult task and engaging stimuli, it can be difficult to predict where humans will look in a complex scene. However, the combination of a difficult search task - namely, breaking the camouflage of a cuttlefish in a scene - and the visually engaging textures in an underwater scene can significantly increase the predictability of the visual cortical signals of subjects [38, 39, 40, 41, 42, 43], from which one can infer increased predictability of gaze. As such, this investigation will also contribute to the rich corpus of research related to predicting fixations of human observers during a search task. Since the human subjects knew what a cuttlefish looks like prior to the visual search task, I hypothesize that a human subject searching for a cephalopod in a given scene has some initial prediction of where a cuttlefish is most likely to camouflage in that scene, i.e. which textures in the scene the cuttlefish is most likely to attempt to match. Furthermore, due to human skill at visual search and pattern recognition, I hypothesize that that initial prediction is most likely correct, regardless of whether the subject ends up identifying the cuttlefish within the trial time limit. In other words, I hypothesize that any agent searching for a cephalopod in a scene will search the regions where the texture distance between cephalopod and background is low.

3.3 Methods

Precision human eye-tracking was performed by Dr. Leslie Osborne and graduate student Neerav Goswami. Figure 3.1 outlines the general structure of the search trial for these subjects, and also shows a composite heatmap of eye positions for a particular image across subjects. The following description of the search trial was provided by Neerav Goswami: "First, fixations for each subject were identified using the EyeMMV Toolbox. Gaze locations were converted to fractions of the total screen resolution as the images had different absolute sizes on the two setups. A fixation was defined to be a series of gaze locations that were within approximately 110 pixels for the 1920x1080 monitor and 147 pixels for the 2560x1440 monitor. Additionally, the gaze locations must have remained in that area for at least 50 ms. The first fixation following the center spot was discarded. Each fixation location was convolved with a 2D Gaussian with a standard deviation of $sigma = 2$ degrees of visual angle. The Gaussians were then scaled by the amount of time the fixation took.

$$C(x,y) * G(x,y) = \sum_{i=1}^{N_{fix}} \frac{1}{\tau} C(x,y) * (\frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}})$$

Where $C(x,y)$ is the location of the center of a fixation that lasts τ milliseconds. Composite heat maps were created by summing the individual subject heat maps and were normalized by dividing them by the peak value." The image set used consist of 105 images, and were collected by Dr. Roger Hanlon and his team in Turkey, and depicts the common cuttlefish *Sepia officinalis* camouflaging in various reef scenes. The fixation data for 19 human subjects were collected for each of these images. While these composite heat maps are a helpful illustration of the overall tendencies of the average searcher, the individual fixation data for each subject was also separately logged in a data structure, and many of the following analyses require resolution on the scale of individual subjects rather than a composite. Armed with this fixation data, an easy first step is to find out whether the fixation patterns

of an individual subject are predictable.

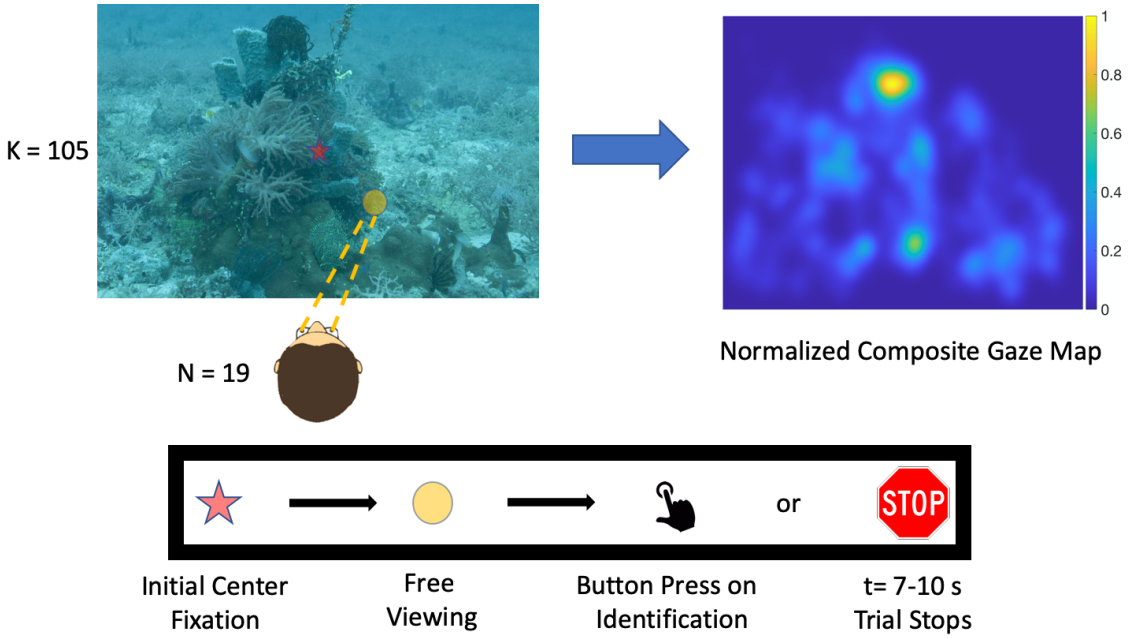


Figure 3.1: Trials were conducted on human subjects (N = 19) in which they searched for and attempted to positively identify cuttlefish in a set of natural scene images (K = 105). Precision eye-tracking gives the location of visual fixations in a given scene, as shown in the across-subject composite for the example image.

3.4 Results

The hypothesis - that any agent searching for a cephalopod will search the regions where the texture distance between cephalopod and background is low - can be tested in a number of ways. A helpful and easy first step is to ask whether the texture distance heatmaps correlate with the human gaze dwellmaps. When calculating the heatmaps, all the values within two times the estimated body radius from the center of the cuttlefish were set to not-a-number. This not-a-number setting is maintained throughout the analyses detailed in this chapter. This kind of conservative masking allows us to ask whether a human observer has some prior prediction of the textures that a cephalopod is most likely to match in a scene. If a fixation inside or immediately outside the cephalopod were considered in the

calculation, we run the risk of comparing a region where the human fixated to the cuttlefish itself, which would, by definition, have a low texture distance value. So, by masking out the pixels inside and immediately in the vicinity of the cephalopod, the comparison is more conservative. If, for a collection of fixations that are completely spatially isolated from the cephalopod, the texture distance values at those locations are still low, then it is safe to assume that the observer had some mental template or prediction of the textures the cephalopod is most likely to match in the scene. The not-a-cephalopod patch (explained below) is also masked out so that similar biasing does not occur in the control case, but this is probably not as critical a step and could be skipped in future analyses. Of course, all this must be regarded as an initial step, and by no means a restriction on other masking possibilities in future experiments. It is possible that fixations close to, but not inside of the cuttlefish contain interesting information about human search strategies, and tighter masking may be employed in the future to investigate this. The texture distance heatmaps and the human gaze dwell maps have a -11.37% correlation on average. The negative sign of the correlation suggests an anti-correlation - broadly speaking, where the human subjects spend more time looking, the texture distance heatmaps are more likely to be low. While this suggests support for our hypothesis (that any agent searching for a cephalopod will search the regions where the texture distance between cephalopod and background is low), the magnitude of the correlation on its own is simply too small to conclusively confirm the hypothesis. Therefore, more sophisticated analyses are necessary.

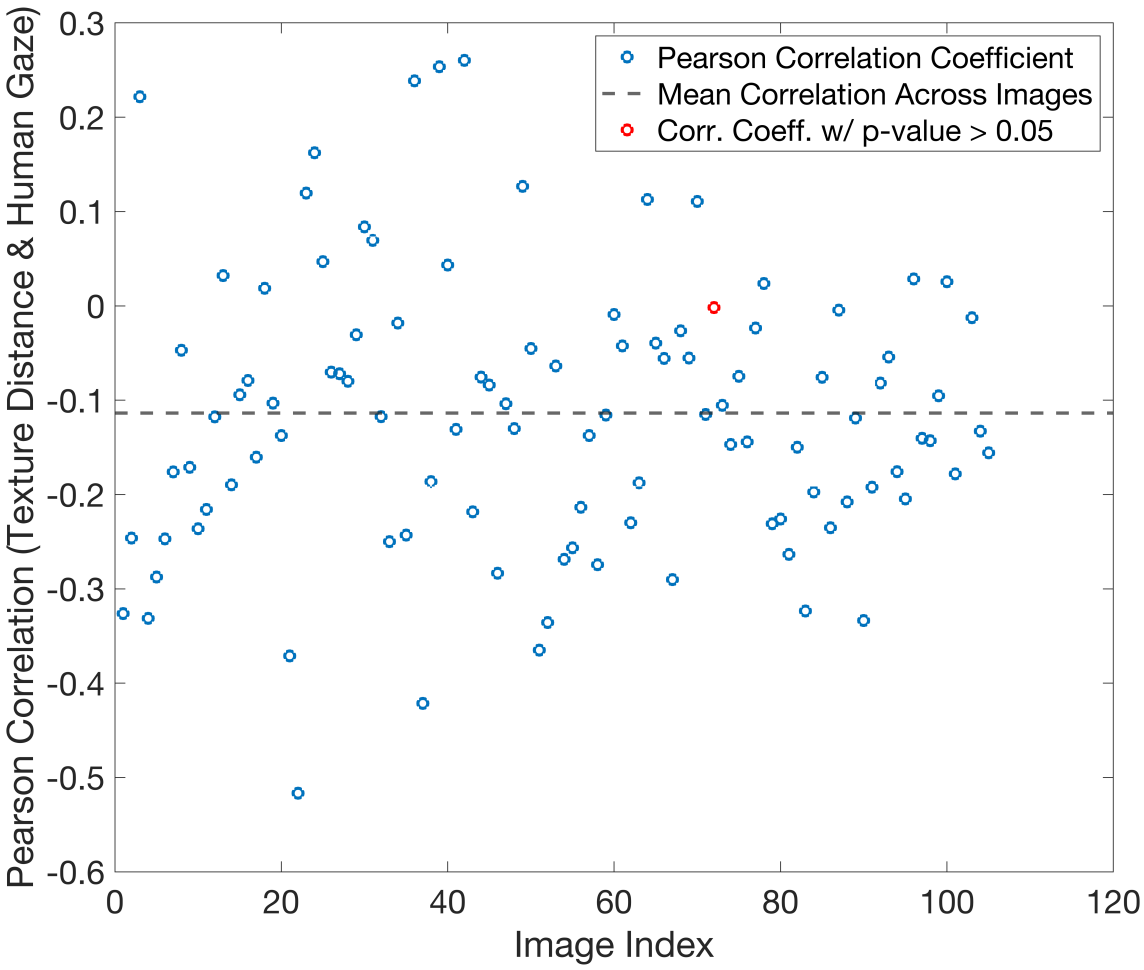


Figure 3.2: The texture distance heatmaps have an (on average) -11.37% correlation with the human gaze dwellmaps. Correlations with p values greater than 0.05 are marked in red. The vast majority of correlations here are statistically significant.

Since it is hypothesized that any agent searching for a cephalopod will search the regions where the texture distance between cephalopod and background is low, a helpful next step would be to investigate how the proportion of fixation locations behaves as the texture distance approaches zero. In this way, it can be clearly established what is meant by "low." Performing these calculations for successful identifications and unsuccessful identifications separately can test the second part of the hypothesis, namely that those who successfully identify the cephalopod will spend a higher proportion of time in the low-texture distance regions than those who failed. Of course, it is critical to control for the choice of reference

patch as a sanity check in order to test the hypothesis - the hypothesis holds only if the texture distance between cephalopod and background is lower for successful agents. If, however, the texture distance between a patch from a region the cephalopod is not trying to match and the background is indistinguishable from the texture distance taken with reference to the cephalopod, then one cannot claim the texture distance metric is a reliable predictor of human gaze. Therefore, for a given image, it is necessary to generate heatmaps where the texture MSE is calculated with reference to a patch selected from a region that the cephalopod is not attempting to match, in addition to the style loss heatmap calculated with reference to a patch on the cephalopod. These patches were selected by identifying intuitively salient features or objects in a given image, then confirming with the cephalopod-reference heatmap that the region selected has a high texture distance from the cephalopod. From this point forward the control context will be referred to as the "not-a-cephalopod context."

So, in both cephalopod and not-a-cephalopod contexts, the proportion of subjects' fixation locations at which the texture distance value is lower than a threshold value is calculated, scanning a range of threshold texture distance values from zero to the maximum. For each image, two composites are made: (1) of the texture distance values at fixation across subjects who failed to identify the cephalopod, and (2) of the texture distance values at fixation across subjects who succeeded in identifying the cephalopod. Each of these composites can then be thresholded by calculating the proportion of fixation locations at which the texture distance is lower than a threshold value for failed and for successful subjects, respectively. This process is then repeated in the control context. Each curve in Figure 3.3 gives the average value of this proportion across images. The error bars give the standard error of the value of this proportion across images, and the legend specifies the context. As explained above, the centers of fixations were saved in a data structure, and pixels within a 1° diameter around the center were considered as part of the fixation. Differing texture distance

distributions among images can be accounted for by performing min-max normalization on the texture distance values at fixation for each image using the minimum/maximum texture distance value in the heatmap for that image.

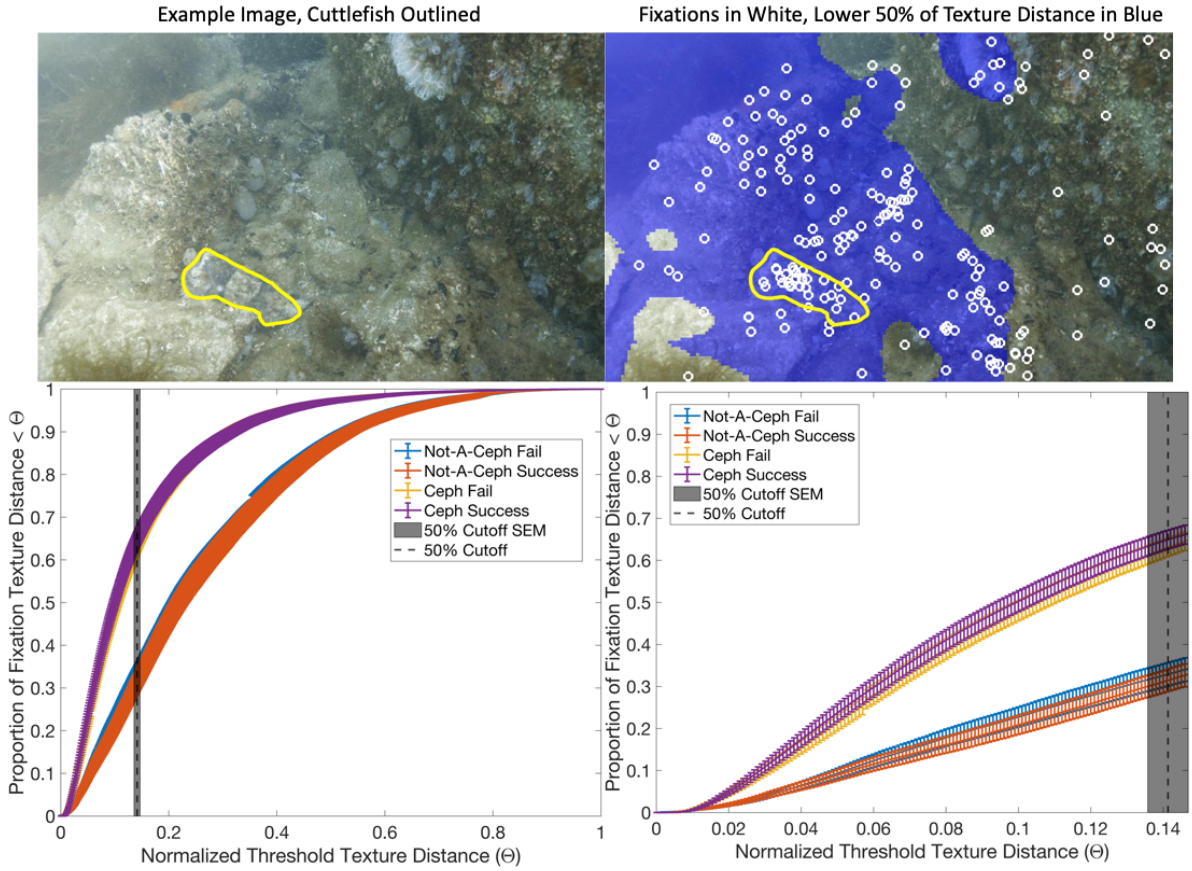


Figure 3.3: (Bottom) Subjects searching for a cephalopod will search the regions where the texture distance between cephalopod and background is low. The majority of texture distance values at fixation in a cuttlefish search task fall below the mean 50% cutoff in the cephalopod context. This does not hold for a control context taken with respect to an obvious non-cephalopod region, so our negative control behaves as expected. (Top) A striking demonstration of the hypothesis in an example image. Images from Turkey by Dr. Roger T. Hanlon and team.

Here, the proportion over the entire domain of texture distance values is plotted, and the "low" texture distance region is more precisely defined. The low region is defined as falling in the bottom 50% of normalized texture distance values. The mean of this 50% cutoff point across the image set is plotted as a black dotted line in Figure 3.3, alongside the standard

error of the cutoff point across the image set. The purple and yellow curves, respectively, show the proportion of texture distance values at fixation locations that are below the normalized threshold texture distance when a reference patch is taken with respect to the cephalopod. The red and blue curves, respectively, show the proportion of texture distance values at fixation locations that are below the threshold when a reference patch is taken with respect to a region that the cephalopod is obviously not trying to match. The bottom left panel shows the behavior of these curves across the normalized texture distance domain, and the bottom right shows the behavior of the curves in the low region. The first and most striking takeaway is the difference between the control curves and the cephalopod context curves. Reading the curves as cumulative distribution functions, the higher values of the cephalopod context curves imply that the texture distance values at fixation in the cephalopod context are lower than the texture distance values at fixation in the control context.

Bootstrapping at the low texture distance cutoff point can be used to determine whether the difference between the control curves and the cephalopod context curves is significant. This method builds the sampling distribution (N = 1000) of the mean proportion of fixation texture distances lower than the cutoff. Comparing these sampling distributions of the mean between the control context and the cephalopod context can reveal whether the difference between the two contexts is significant. If a normalized threshold texture distance of 0.141 is taken as the low texture distance cutoff, the top panel of Figure 3.4 shows the sampling distribution of the mean proportion of fixation texture distances lower than the cutoff for the failed trials in the control context and the failed trials in the cephalopod context. The bottom panel of Figure 3.4 shows the sampling distribution of the mean proportion of fixation texture distances lower than the cutoff for the successful trials in the control context and the successful trials in the cephalopod context. Since there is no overlap between the sampling distributions of the mean proportion between the control context and cephalopod context, the difference between the control curves and cephalopod context curves can be regarded as

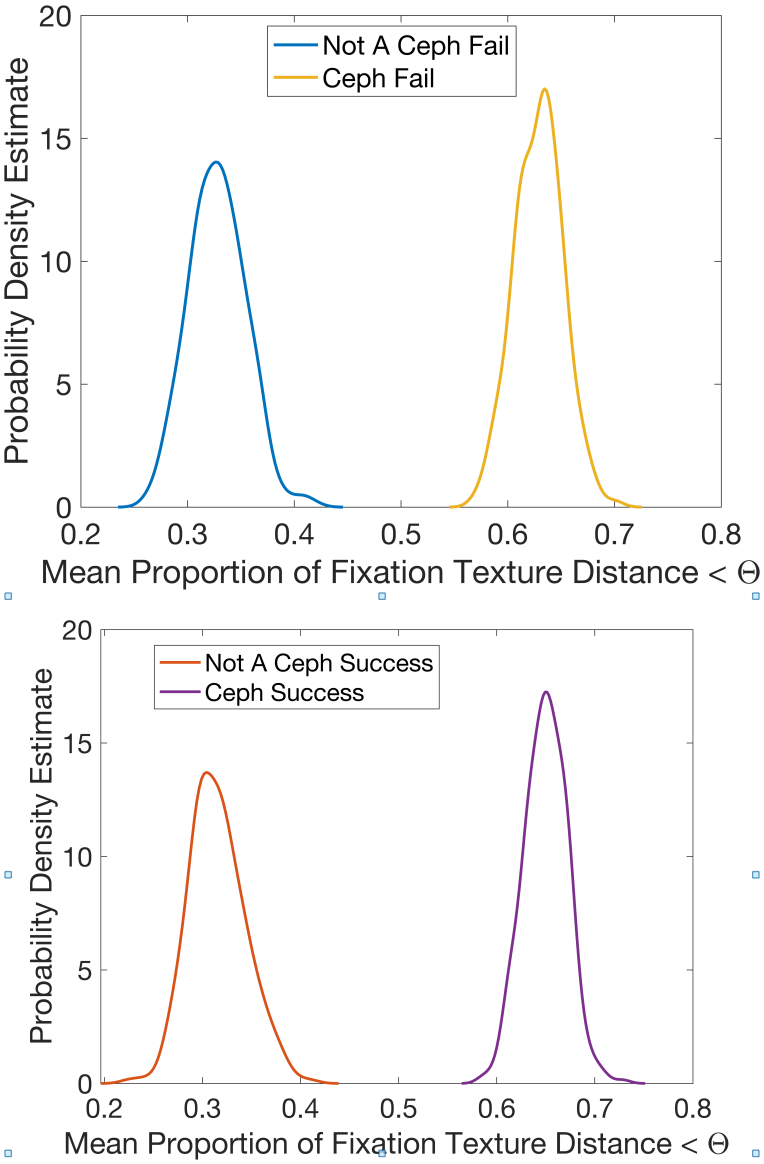


Figure 3.4: Comparing sampling distributions of bootstrapped means of proportion of fixation texture distance below the low texture distance cutoff indicates a statistically significant difference between control curves and cephalopod context curves in both failed/successful trials.

significant. In the low texture distance region, the cephalopod curves comfortably exceed the 0.5 proportion before the lower bound of the 50% cutoff, while the not-a-cephalopod curves remain below the 0.5 proportion for the extent of the low texture distance region. In other words, when a subject searches for a cephalopod, they spend the majority of their fixations searching regions where the texture distance from the cephalopod is low (indicating textural similarity to the cephalopod), while comparatively neglecting regions where the texture distance to an obvious non-cephalopod object is low. This not only shows that our control curves behave as they should, but also that our hypothesis is directly confirmed: any agent searching for a cephalopod will search the regions where the texture distance between cephalopod and background is low. The top panels of Figure 3.3 show a striking demonstration of the hypothesis - note that the vast majority of fixations fall in the blue mask indicating the bottom 50% of normalized texture distance values.

Now, bootstrapping at the low texture distance cutoff point can also be used to examine whether the difference between successful and failed trials within the control context is significant, and whether the difference between successful and failed trials within the cephalopod context is significant. The top panel of Figure 3.5 shows the sampling distribution of the mean proportion of fixation texture distances lower than the cutoff for the failed and successful trials in the control context. The bottom panel of Figure 3.5 shows the sampling distribution of the mean proportion of fixation texture distances lower than the cutoff for the failed and successful trials in the cephalopod context. Here, due to the high level of overlap between the sampling distributions of the mean proportion for failed and successful trials in both contexts, no claim can be made based on the proportion of fixation analysis alone that a significant difference exists between successful and failed subjects, although the general effect of search strategy is significant and supports the hypothesis. In order to more deeply investigate the difference between failed and successful subjects, it may be informative to take into account the duration of each fixation, rather than treating them equally as in the

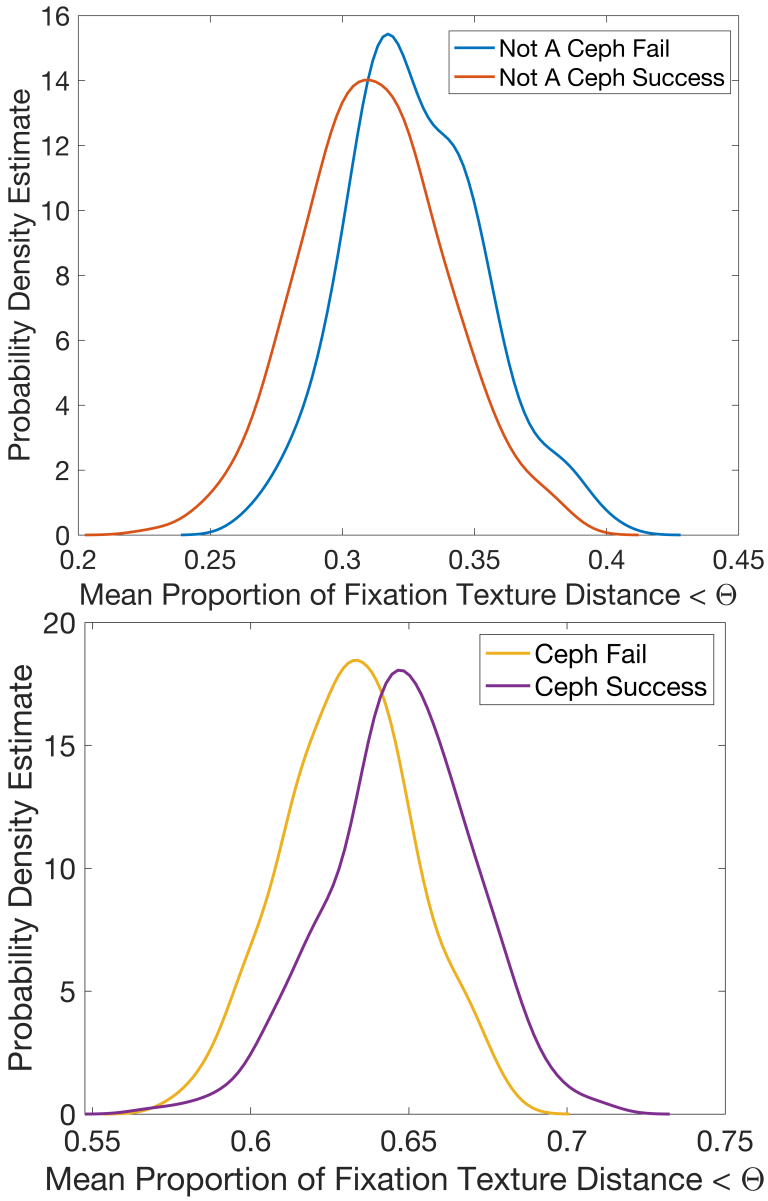


Figure 3.5: Comparing sampling distributions of bootstrapped means of proportion of fixation texture distance below the low texture distance cutoff does not indicate a statistically significant difference between failed and successful trials within control context and cephalopod context.

above analysis.

To better understand how to incorporate duration of fixation into the analysis, consider a single fixation made by a failed subject on a given image, as shown in Figure 3.6. To begin, the proportion of this fixation in which the texture distance values are below a given threshold is calculated. This proportion is then multiplied by the duration of the fixation, effectively adjusting the duration according to the amount of time spent fixating on below-threshold texture distance values. This process is repeated for every fixation for that particular subject, then the resulting adjusted durations are summed. Finally, the sum of adjusted durations is divided by the total trial time. If N_{fix}, t_{fix} is the duration of the fixation, t_{trial} is the total trial time, B_τ is the number of pixels below the threshold τ, and A is the total number of pixels in the fixation area, the above calculation can be represented by

$$\rho = \sum_{i}^{N_{fix}} \frac{t_{fix}(B_\tau / A)}{t_{trial}} \tag{3.1}$$

Next, ρ is averaged across all failed subjects for a single image, giving ρ_{subj}, which in turn is averaged across the images in the dataset, giving ρ_{img}. The standard error of ρ_{img} is given by

$$\sigma_{img} = \frac{1}{N_{img}} \sqrt{\sum \sigma_{subj}^2},$$

where

$$\sigma_{subj} = \frac{std(\rho_{subj})}{\sqrt{N_{subj\_fail}}}$$

and where N_{subj_fail} is the number of subjects who failed to identify the cuttlefish. In order to calculate the same values for successful subjects, we substitute the number of subjecs who succeeded in identifying the cuttlefish $N_{subj\_success}$ for N_{subj_fail}. Finally, the value of ρ_{img} is calculated for the range of normalized threshold texture distance values between 0 and 1. The steps of this process are represented visually and summarized in Figure 3.6.

The plot of ρ_{img} against the normalized threshold texture distance is shown for failed and

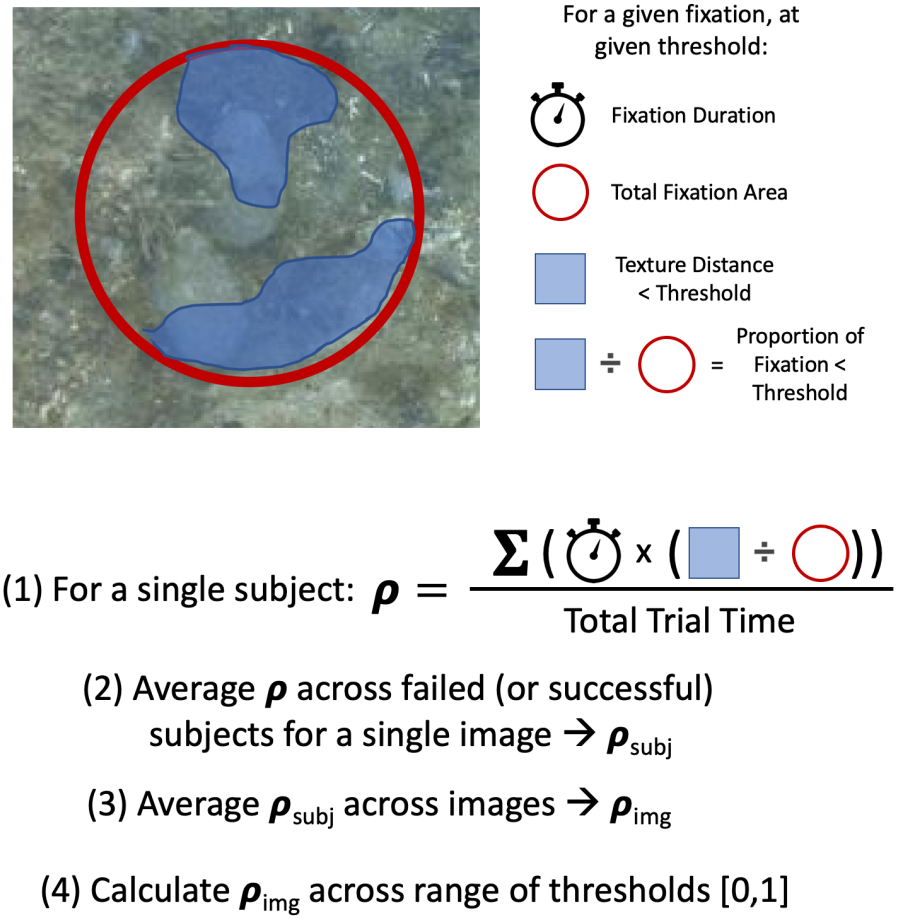


Figure 3.6: Visual illustration of incorporation of duration into threshold calculation. (Note: this does not show a real fixation location or threshold values, this fixation was simulated to demonstrate the process of calculating ρ_{img})

successful trials in both cephalopod and not-a-cephalopod contexts in Figure 3.7. (A helpful note: here it makes sense that the curves do not saturate at 1 as in the above analysis, since subjects do not spend 100% of the trial time fixating).

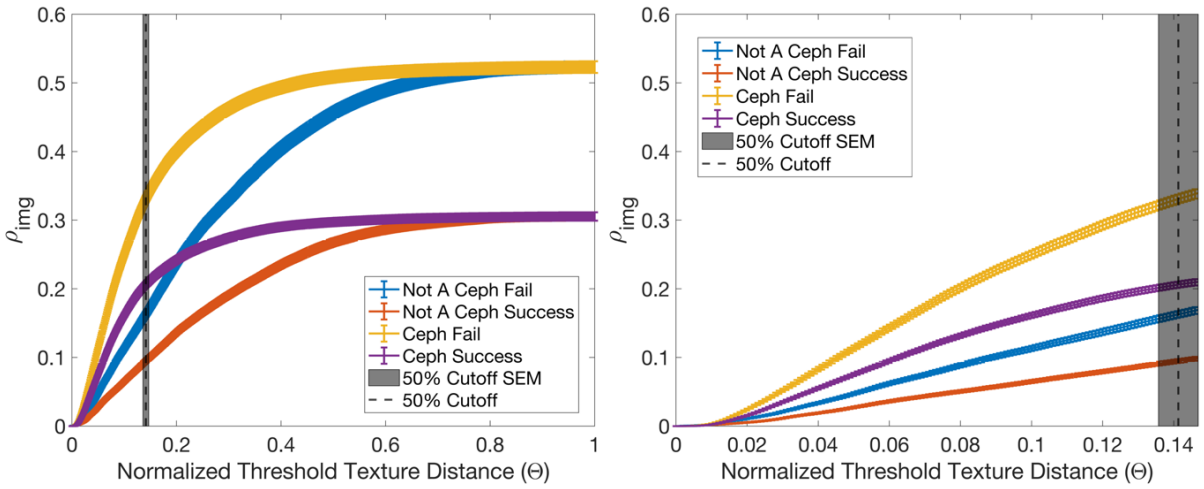


Figure 3.7: Subjects searching for the cephalopod will spend a higher proportion of time fixating in regions that are texturally similar to the cuttlefish compared to regions that are texturally similar to other objects. Subjects who fail to identify the cephalopod spend more time fixating in low texture distance regions than those who succeed.

Here, below the 50% cutoff point, the values of ρ_{img} in the cephalopod context success curves are higher than the ρ_{img} in the control context success curves, and the values of ρ_{img} in the cephalopod context failure curves are higher than the ρ_{img} in the control context failure curves. Qualitatively, higher values of ρ_{img} can be equated with a higher proportion of time spent fixating on values below the normalized threshold texture distance.

As before, the bootstrapping method can be used to determine the significance of the difference between control context and cephalopod context curves. If a normalized threshold texture distance of 0.141 is taken as the low texture distance cutoff, the top panel of Figure 3.8 shows the sampling distribution of the mean of ρ_{img} at the cutoff for the failed trials in the control context and the failed trials in the cephalopod context. The bottom panel of Figure 3.8 shows the sampling distribution (N = 1000) of the mean of ρ_{img} at the cutoff for the successful trials in the control context and the successful trials in the cephalopod context. As

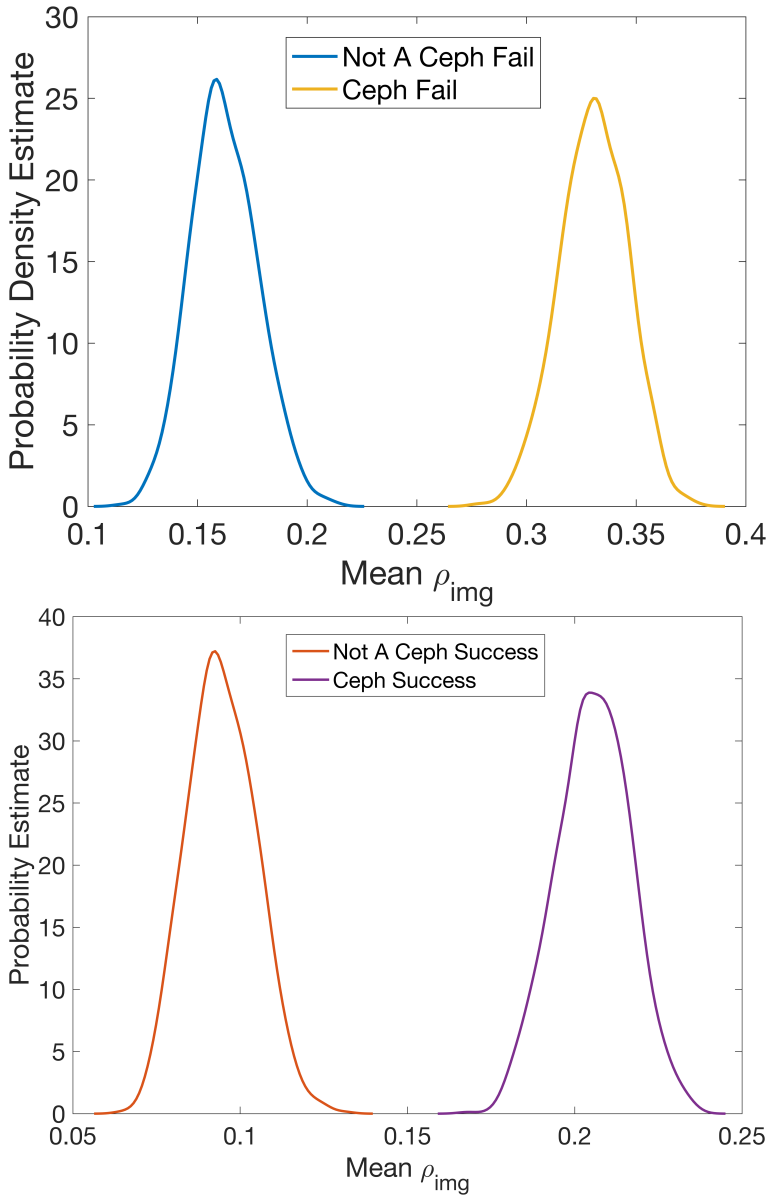


Figure 3.8: Comparing sampling distributions of bootstrapped means of ρ_{img} at the low texture distance cutoff indicates a statistically significant difference between control curves and cephalopod context curves in both failed/successful trials.

before, the lack of overlap suggests the difference between the control curves and cephalopod context curves can be regarded as significant. The values of each of the cephalopod context curves are greater than their corresponding control curves in the low-threshold regime. This suggests that subjects searching for the cephalopod will spend a higher proportion of time fixating in regions that are of low-texture distance from the cuttlefish than in regions that are of low-texture distance from obviously non-cuttlefish objects. In other words, it would make sense for subjects searching for a cuttlefish to spend an higher proportion of time in regions that are texturally similar to cuttlefish than to other objects. This means the negative control once again behaves as expected.

Once again, bootstrapping at the low texture distance cutoff point can also be used to examine whether the difference between successful and failed trials within the control context is significant, and whether the difference between successful and failed trials within the cephalopod context is significant. The top panel of Figure 3.9 shows the sampling distribution of the mean of ρ_{img} at the cutoff for the failed and successful trials in the control context. The bottom panel of Figure 3.9 shows the sampling distribution of the mean of ρ_{img} at the cutoff for the failed and successful trials in the cephalopod context. The key difference from the previous analysis is that, since the sampling distributions have no overlap, the separation between successful and failed subjects in the cephalopod context is now statistically significant. The domination of the failure curve in the cephalopod context implies that subjects who fail to identify the cephalopod will spend a higher proportion of time fixating in low texture distance regions than those who succeed in identifying the cephalopod. In other words, subjects who succeeded needed to spend less time fixating in low texture distance regions before they were able to identify the cephalopod. This result would also be consistent with the frustration of a subject in a failed trial as trial time progresses. This frustration might lead to more random search patterns, a higher frequency of fixations, and changing search strategies. This would suggest that ρ_{img} could be used as a measure of

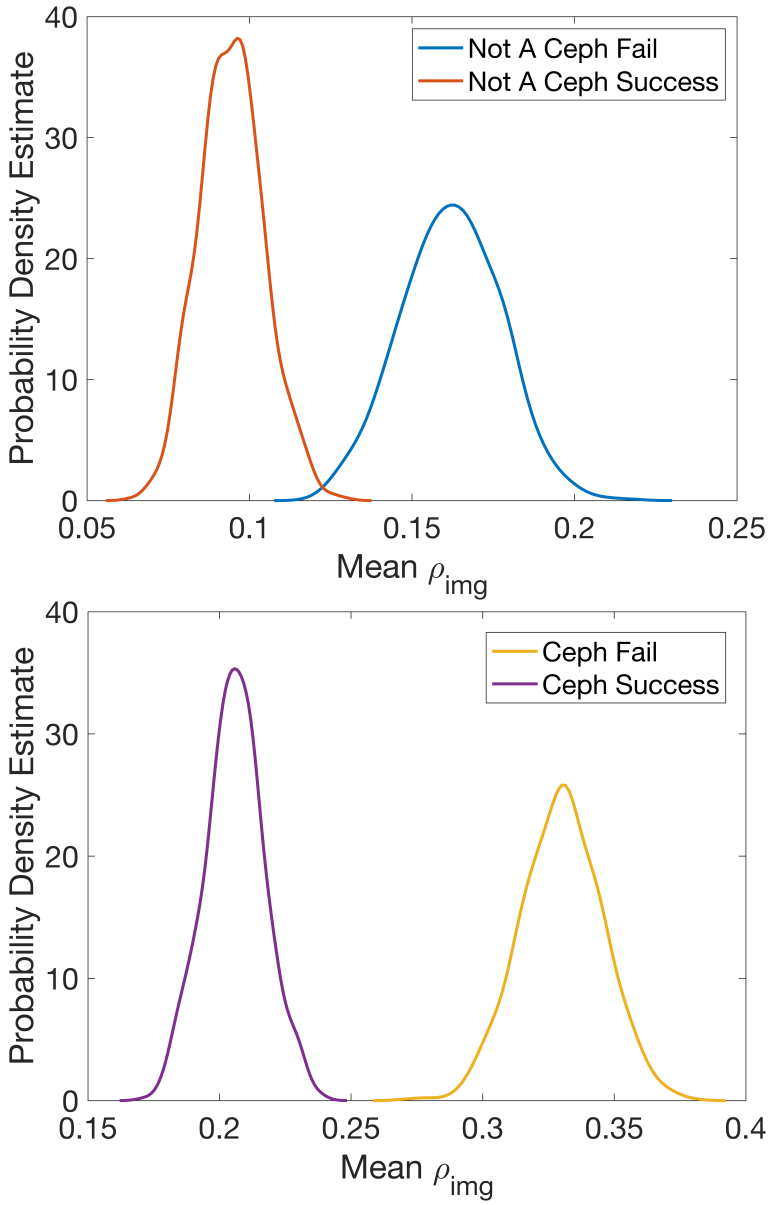


Figure 3.9: Comparing sampling distributions of bootstrapped means of ρ_{img} at the low texture distance cutoff indicates a statistically significant difference between failed and successful trials in the cephalopod context.

a subject's skill at identification.

It is also worth asking why a separation exists in the failure and success curves for the control context, although the slight overlap in the sampling distributions suggests that this separation is less statistically significant than in the cephalopod context. Note that this is not incompatible with the hypothesis, since both control curves fall below the cephalopod context curves in the low texture distance regime, but the separation is still curious considering the trial does not involve searching for an obviously non-cephalopod object. This is a ripe area for future investigation, but perhaps this separation speaks to the possibility that most objects in an underwater scene restricted to a single ecological niche resemble each other, i.e. the "obviously different" non-cephalopod objects are simply the furthest away on that scale, but still more similar to a cuttlefish than, say, a picture of an object in a completely unrelated above-ground scene. In any case, the separation between the success and failure curves in the cephalopod contexts shows that the texture distance metric not only identifies the general search strategy of human subjects, but is capable of distinguishing between successful and failed subjects.

3.5 Discussion

The texture distance metric developed in Chapter 2 was validated and tested through the use of precision human eye-tracking data. It was hypothesized that subjects searching for the cuttlefish will search regions of the scene where the texture distance between cuttlefish and background is low. The eye-tracking experiments establish that the texture distance metric is indeed able to predict the general search strategy of human subjects, regardless of whether the search trial is successful or failed. That is, subjects will search regions that are texturally similar to the camouflaged cuttlefish in the scene, indicating that human subjects have some idea of what a cuttlefish in a given scene is likely to look like. It was found that the texture metric is also capable of distinguishing between successful and failed trials

after accounting for the duration of visual fixations. Specifically, it was found that subjects who fail to identify the cuttlefish spend a higher proportion of time searching regions of low texture distance than those who succeeded. This might suggest that subjects who succeed need to spend less time in low texture distance regions before finding the cue that led to the identification of the cuttlefish. It may also suggest that the frustration of subjects in failed trials leads to changing search strategies mid-trial or more random fixation patterns. Confirming or rejecting these suggestions is a rich area for future research. A potential approach to this problem is to characterize the last fixation locations made during successful and failed trials.

Additionally, in the analysis presented above, a circle of two estimated body radii centered on the cuttlefish was disregarded, as well as the location of the reference patch in the control, not-a-cephalopod context. Reiterating these analyses with less conservative masking - for instance masking with tighter ellipses or freeform shapes rather than circles - may also reveal more information about predator search strategies.

The existence of a relationship between the texture distance metric and the time spent fixating by human visual searchers is significant, and forms the basis of a promising method for camouflage quantification. In other words, it is safe to assume that the calculated values of the texture distance metric in a scene are not disconnected from the measured reality of the average predator's visual search of that scene. It therefore follows that, given only the histogram of texture distance values in a scene, it should be possible to read out the difficulty of the search task, i.e. the texture distance metric is validated as a promising avenue for camouflage quantification. However, the ability to rank the effectiveness of camouflage in a scene given *only* the histogram of texture distance values should be the immediate focus of future research. An easy initial step would be to separate the image bank into two categories based on low and high rates of correct cephalopod identification. Identifying common features of the texture distance distributions within each group may point toward

a method for ranking difficulty or the effectiveness of camouflage in a scene.

In addition to validating the texture distance metric as a promising method for camouflage quantification, the above results also contribute to existing research on the challenging problem of predicting human gaze, so it is important to place the results in the context of previous studies in this field. The greater proportion of time spent in low texture distance regions is, for instance, reminiscent of the results of the human gaze research of Rajashekar, et al. [44], neatly summarized in Geisler, et al. [45] as "the eye is attracted (at least some of the time) to features in the noise that match features of the target." In this study, the authors [44] present human observers with a search task of a target embedded in 1/f noise, and then extract a discrimination image from the noise to identify the features that capture the attention of the observers. In the case of a cephalopod search task, the "noise" is admittedly more complicated than the 1/f noise used in the Rajashekar, et al. [44] study, since it encompasses any part of an underwater scene that is not the target cephalopod, but the analogy is still instructive. The higher proportion of time spent in low texture distance regions during the cephalopod search task similarly conveys an interest in "features in the noise that match features of the target", but does so in a much higher dimensional context and exploits the power of deep neural networks, whose sophistication has grown significantly since the Rajashekar, et al. [44] study was published. One can also connect the results of this chapter to a possibility suggested in the Discussion of the Rajashekar, et al. [44] study, namely that the discrimination images extracted could be used as a kernel to predict where human observers would likely fixate. Similarly, the results show that the reference patch on the cuttlefish, used in the context of a texture synthesis algorithm, can be used as a kernel to predict where human observers will likely look, i.e. the low texture distance regions of the texture distance heatmaps. As such, it is reasonable to view the results of this chapter as a significant reinforcement of the conclusion in Rajashekar, et al. [44] and therefore a contribution to the current state of human gaze research.

The results of this chapter also add a new perspective to current research regarding predation on cephalopods, specifically. The responses of a range of potential predators to cuttlefish was observed in a laboratory setting by Staudinger, et al. [16], and predation has been observed in the wild [4, 5, 3]. There have been few perspectives on predation that have been *first-person*, however, and the results of this chapter serve to address this scarcity. Although the predators used are human, the results still give insight into which regions a scanning predator may be most likely to search.

While the possibilities for future research are boundless, the results presented above point to a promising method for camouflage quantification, and additionally give valuable insight into the prediction of human gaze. The ability to understand not only where humans generally look, but also where they look when they succeed/fail at identification during a search trial, would be extremely valuable in the context of devising effective artificial camouflage methods and technologies.

CHAPTER 4

INVESTIGATING COLOR AND LUMINANCE DISCRIMINABILITY IN PREDATOR VISUAL SPACE VIA HYPERSPECTRAL IMAGING

This work was completed in collaboration with Dr. Roger Hanlon, Dr. Derya Akkaynak and Dr. Kendra Buresch, who collected and organized the hyperspectral images used in this study, Dr. C.C. Chiao, who identified which days the data from the camera was usable, and Dr. Stephanie Palmer.

4.1 Abstract

Here, I use underwater hyperspectral images (HSI) to simulate the predator views of natural scenes of cephalopod camouflage. Choosing three natural predators of cuttlefish varying in number of photoreceptors and level of predation on the common cuttlefish *Sepia officinalis*, I investigate their ability to chromatically discriminate the cephalopod from its background using a receptor noise-limited model for color vision. I also examine their ability to discriminate the cephalopod from its background based on luminance alone. The results of this investigation suggest that cephalopods appear to have evolved in a generalist rather than a specific way, i.e. to be minimally discriminable regardless of the number of photoreceptors in the predator visual system. However, in all predator cases, camouflaged cephalopods are more discriminable through differences in luminance than through differences in color. This result is consistent with the conclusions of previous HSI work on camouflage, but this study uses real rather than hypothetical predator visual systems for simulation.

4.2 Introduction

Cephalopod camouflage strategies were shaped in order to evade detection and recognition by (primarily teleost) predators in the ecological niche that cephalopods inhabit. Research regarding the quantification of cuttlefish camouflage would benefit immensely from a strategy that allows researchers to "look through the eyes" of these natural predators.

In fact, the use of hyperspectral imaging (HSI) has been proposed and tested for this purpose. Each pixel in a hyperspectral image gives the reflectance spectrum across a number of wavelength bands of the space represented by that pixel. This can be contrasted with the relatively simple structure of a standard RGB image, an approximation of human vision based on the three wavelength bands corresponding to the peaks of human visual pigments. The abundance of spectral bands, extending from ultraviolet to infrared wavelengths, turns each image into an extremely rich dataset. This has enabled its use in medical imaging for tumor assessment and dermatology [46, 47], but critically, also allows for the simulation of the vision of any organism with a color-opponent rod/cone-mediated visual system [2]. Provided the spectral sensitivity of a model organism is known, taking the dot product of each cone spectral sensitivity curve and the reflectance spectrum at each pixel of a hyperspectral image approximates the cone signal generated by the represented scene in the organism's visual system. This cone signal can in turn be used to measure the chromatic discriminability of objects in the scene from the perspective of the organism. This simulation capability has been used to investigate predator perception in teleost fish predators in the landmark study by Chiao, et al. [2]. In this latter study, hyperspectral images of camouflaged cuttlefish in a laboratory setting indicated a good color match of the cuttlefish to the background through the eyes of constructed, hypothetical fish predators. This thesis aims to build on this study by reiterating this color-match analysis (1) through the eyes of *real* fish predators, (2) using hyperspectral images of cuttlefish camouflaging in the wild *under natural lighting conditions*. In the case of the hyperspectral images considered in this thesis, this reflectance

spectrum has 16 wavelength bands, and a spatial resolution of 512x512 pixels, such that a hyperspectral image can be visualized as a "data-cube". The mathematics and finer details of this color-match analysis are given along with the results in the following section.

4.3 Methods

The hyperspectral images used in this study were collected in 2016 in the Indonesian archipelago of Raja Ampat, an extremely rich marine habitat. The collection was performed by a team led by Dr. Roger Hanlon at the Marine Biological Laboratory. The collection process consisted of 24 dives across 10 days of fieldwork, with the team working virtually around-the-clock for the extent of the trip. Hyperspectral images were collected with a Surface Optics SOC Model #716-UVA. A day-by-day analysis of images taken of the Spectralon white-standard during the trip was conducted by Dr. C.C. Chiao - this analysis determined that the filter tray of the hyperspectral camera had been shifting throughout the trip, such that certain days would require further forensic re-calibration. However, on the basis of the Spectralon analysis, images taken on Days 2, 4, 9, & 10 of the trip were determined to be usable as is, with the caveat that data from the eighth spectral band (500 nm) be discarded. Since this thesis is concerned with background-matching camouflage, those images of cuttlefish that were obviously in motion or stark against the water column were disregarded. The analysis described in the following section was performed on this truncated dataset (N=43 images). A number of example images, rendered as RGB images, are shown in the Figure 4.1. For one of these hyperspectral images, the visual data in each color channel is also shown as a greyscale image in Figure 4.2

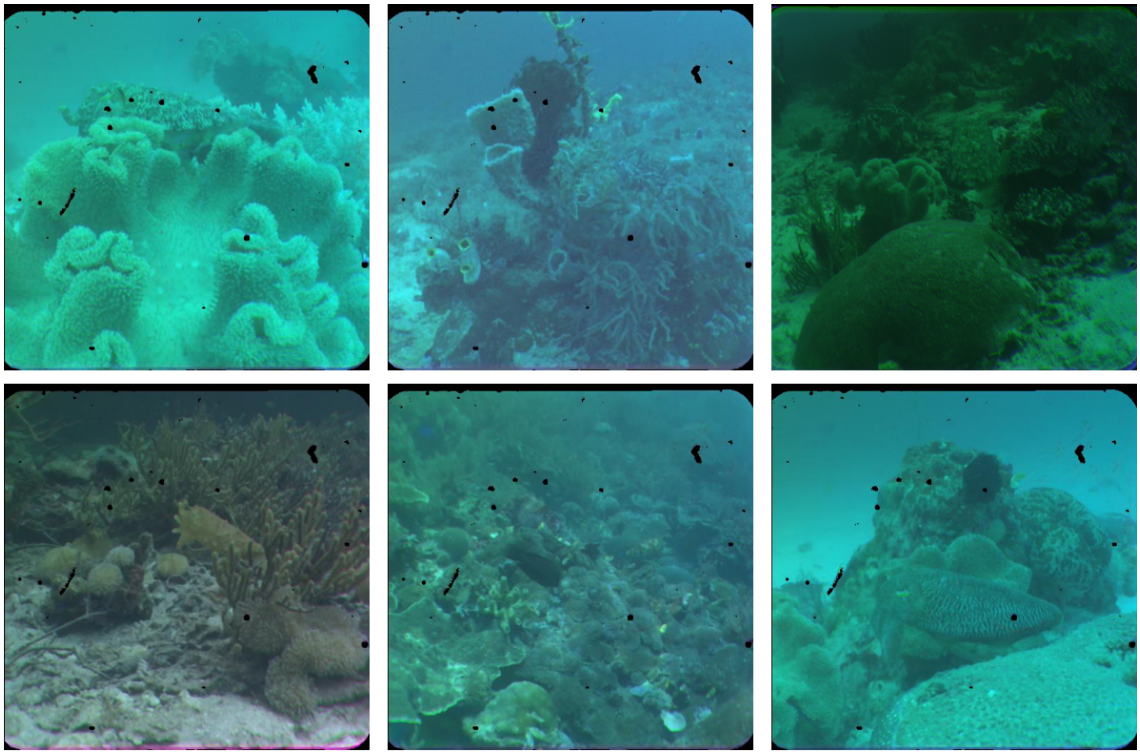


Figure 4.1: Example images taken by the Drs. Roger Hanlon and Derya Akkaynak at Raja Ampat, visualized as RGB images. Note that the actual HSI format is a 512x512x16 data cube, and these RGB renderings are created for viewing convenience.

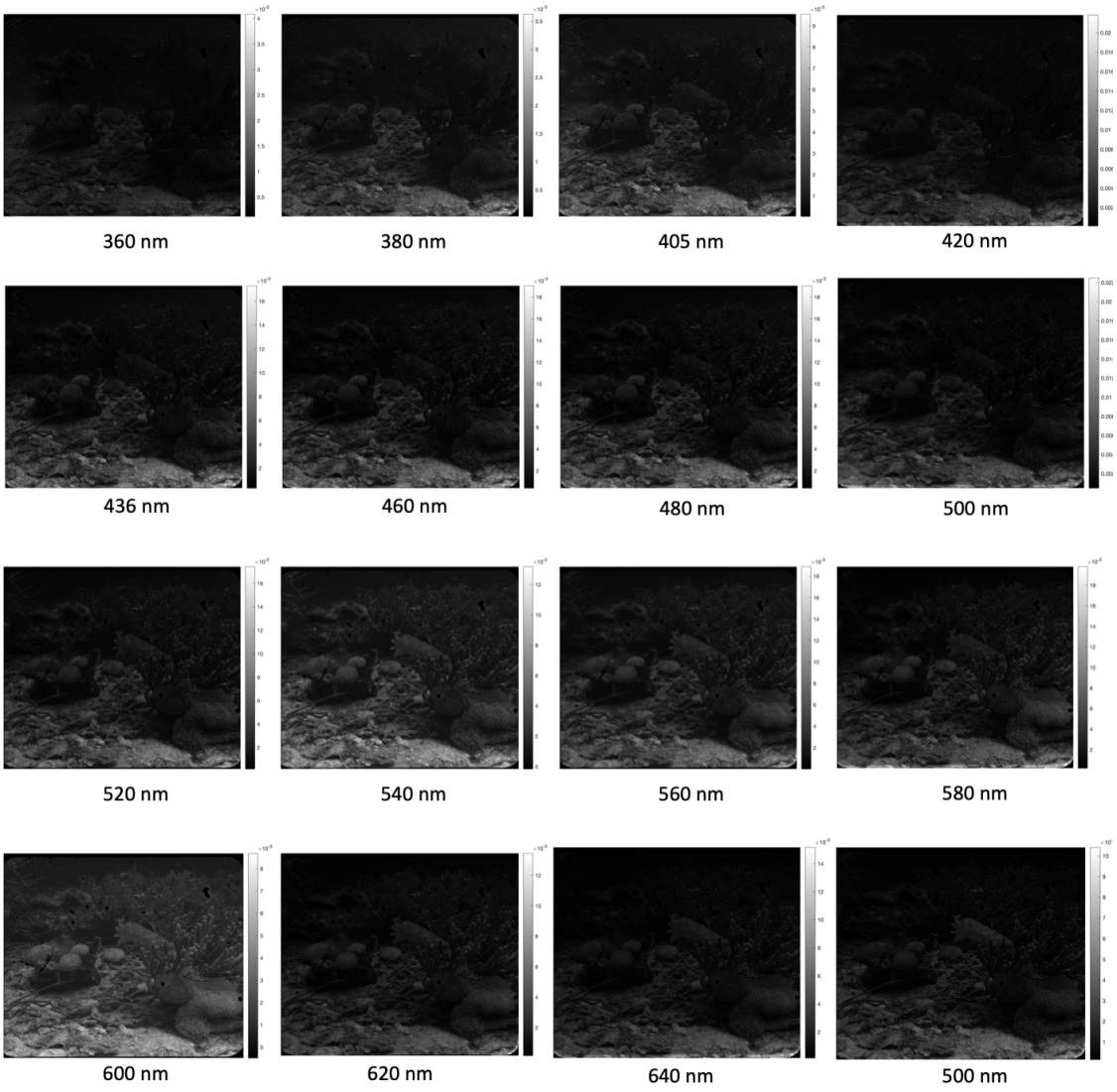


Figure 4.2: An HSI image decomposed across the sixteen color channels, visual data expressed in grayscale.

The authors [2] modeled predator vision by using the Vorobyev and Osorio [48] receptor space model, which prescribes formulae for the threshold on the just-noticeable difference ΔS between two color stimuli for a given visual system. The model relies on a number of core postulates. The threshold distance is limited by noise in photoreceptors, and for n types of photoreceptors, color perception will be governed by distances in an $n-1$ dimensional space. Since any pair of receptor types could constitute a possible opponency, and since ΔS is not limited by the choice of these pairs, all $\binom{n}{2}$ pairwise receptor combinations are accounted for in the formulation of ΔS. The noise in photoreceptors is assumed to be Gaussian, since this model is defined in terms of the standard deviation of noise. Additionally, a unique feature of the model is that there is no opponency mechanism for achromatic stimuli, so luminance is effectively ignored, hence the $n-1$ dimensionality. That is, the model gives zero response if a stimulus and its background differ in luminance, but not in hue. As such, the model performs optimally for scenes under high illumination with large, stationary stimuli, in which the role of the achromatic channel is minimized. The threshold ΔS between two color stimuli relies on the noise in each photoreceptor channel and the quantum catches of each type of photoreceptor for the two stimuli. Since this analysis focuses only on chromatic discrimination, all photoreceptors measured will be cones, which govern color vision in fish. For a given photoreceptor of type i, the difference in sensory response for two color stimuli a and b is given by the difference in the logarithm of the receptor quantum catch for these stimuli. If the quantum catch is given by q_i, $Q_i = \ln(q_i)$, and standard deviation of noise in a receptor of type i is given by e_i, then

$$(\Delta S)^2 = \frac{1}{\sum_{k=1}^{N-1}(\frac{1}{e_k^2})} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{(\Delta Q_i - \Delta Q_j)^2}{e_i^2 e_j^2} \tag{4.1}$$

Here, N represents the number of photoreceptors involved in color perception in a given visual system. To model a dichromatic predator, $N = 2$, to model a trichromatic predator, $N = 3$, and so on. Given a hyperspectral image of a cuttlefish on a natural background and

set of spectral sensitivites representing a predator visual system, this formulation can be used to generate just-noticeable difference (JND) images by calculating the values of ΔS between each pixel and the average of the pixels in the natural background. This gives the probability that the given predator can *just* notice the chromatic difference between the selected pixel and the background. A set of spectral sensitivity curves for a trichromat predator can be given by $S_1(\lambda)$, $S_2(\lambda)$, and $S_3(\lambda)$, where λ is wavelength in nanometers. The reflectance spectrum at a given pixel in a hyperspectral image is given by $R(\lambda)$. Therefore, given a color stimulus, the quantum catch q_i for each photoreceptor can be expressed as

$$q_1 = \sum_{i=\lambda_0}^{\lambda_f} R(\lambda) S_1(\lambda) \tag{4.2}$$

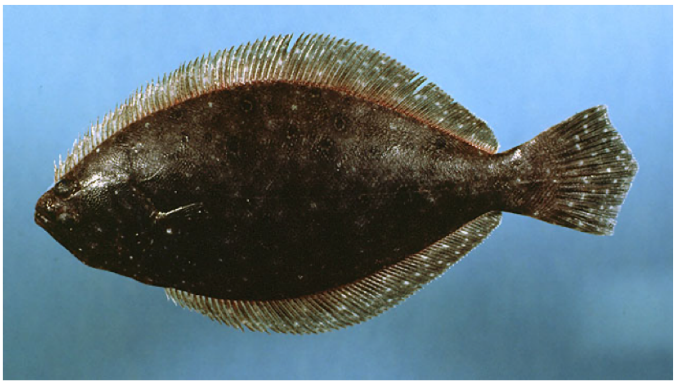
$$q_2 = \sum_{i=\lambda_0}^{\lambda_f} R(\lambda) S_2(\lambda) \tag{4.3}$$

$$q_3 = \sum_{i=\lambda_0}^{\lambda_f} R(\lambda) S_3(\lambda) \tag{4.4}$$

In this analysis, the quantum catches are mean-adapted, and black and white surfaces are explicitly defined in the code to normalize q_i to a range of $[0, 1]$.

European pollock (*Pollachius pollachius*)

European seabass (*Dicentrarchus labrax*)



Summer flounder (Paralichthys dentatus)

Figure 4.3: Three teleost predators selected for HSI analysis, images obtained from FishBase.

European pollock (*P. pollachius*)	Trichromat (458, 498, 521 nm)	Exerts predatory pressure on *S. officinalis*
European seabass (*D. labrax*)	Trichromat (380, 497, 598 nm)	Accessory predation on *S. officinalis*
Summer flounder (*P. dentatus*)	Dichromat (449, 525 nm)	Does not share habitat with *S. officinalis*

Table 4.1: Summary of predators selected for HSI analysis. These predators represent a range of visual system types and levels of predation.

While the $R(\lambda)$ at each pixel is given by hyperspectral imaging, the set of $S(\lambda)$ curves depends on the choice of visual predator. Spectral sensitivity curves give the absorbance of visual pigments as a function of wavelength, and determine the color discrimination abilities of that animal. Which predators should be considered? The species of cuttlefish represented in the HSI dataset are *Sepia latimanus* and *Sepia apama*, which inhabit shallow waters and coral reefs north of Australia [49, 50]. However, the spectral sensitivities of underwater predators of these species were not as readily available as those for the underwater predators of the common cuttlefish (*Sepia officinalis*). The common cuttlefish is native to the eastern North Atlantic and the Mediterranean, and lives in shallow waters on "sandy or muddy substrates" [51]. Despite the difference in habitats, all three species of cuttlefish have impressive cryptic abilities, so it is assumed that a frequent predator of the common cuttlefish would still be well-equipped to hunt *S. latimanus* and *S. apama*. Three real fish predators of the common cuttlefish are considered, in descending order of the frequency of predation: the European pollock (*Pollachius pollachius*) [52], the European seabass (*Dicentrarchus labrax*)[52], and the summer flounder (*Paralichthys dentatus*). The European pollock is a trichromat, and is known to exert significant predatory pressure on (*Sepia officinalis*) in the wild in the Bay of Biscay ($\lambda_{max} = 458, 498, 521nm$ [52], Pigment weights $= 0.8214, 0.3214, 1$). The European seabass is also a trichromat, and also shares an ecological niche with cuttlefish, although cuttlefish are an accessory prey ($\lambda_{max} = 380, 497, 598nm$, Pigment weights $= 0.17, 0.32, 1$)[52]. The summer flounder is a dichromat, and would not

encounter cuttlefish in the wild ($\lambda_{max} = 449, 525$, Pigment weights $= 0.83, 1$). The spectral sensitivities, appropriate templates [53, 54], and relative proportions of pigments for the European seabass and summer flounder was derived in a study by Horodysky, et al. [55] The maximum spectral sensitivity wavelengths for the European pollock were derived in a study by Shand, et al. [56], and were plugged into the GFRKD rhodopsin template used in the Horodysky, et al. paper in order to approximate the full curves [55, 54]. The relative proportions of pigments for the European pollock were calculated based on the number of cells of each type examined in an immature adult specimen in the Shand, et al. study [56]. Images of these predators (obtained from FishBase) are shown in Figure 4.3, their features are summarized in Table 4.1, and the spectral sensitivity curves for these three fish are shown in Figure 4.4.

4.4 Results

Armed with these real spectral sensitivity curves and a rich underwater HSI dataset, the analysis can proceed. Consider a single image of cuttlefish camouflage. The background is defined as a one animal-size surround centered on the cuttlefish. First, the distribution of luminance inside the animal is calculated, and then compared to the distribution of luminance in the background, giving the level of discriminability-via-contrast through the eyes of a given predator. It is assumed that the long-wavelength (L) cone in trichromats, and the medium-wavelength (M) cone in dichromats, is the luminance channel. This assumption is supported by a study of reef fish visual systems conducted by Siebeck, et al. To compare the two luminance distributions, the Kullback-Leibler divergence (D_{KL}) [57] is used. For two probability distributions $f_1(x)$ and $f_2(x)$ over the same set of outcomes, the D_{KL} is given by

$$D_{KL}(f_1||f_2) = \sum_x f_1(x) \log(\frac{f_1}{f_2}) \tag{4.5}$$

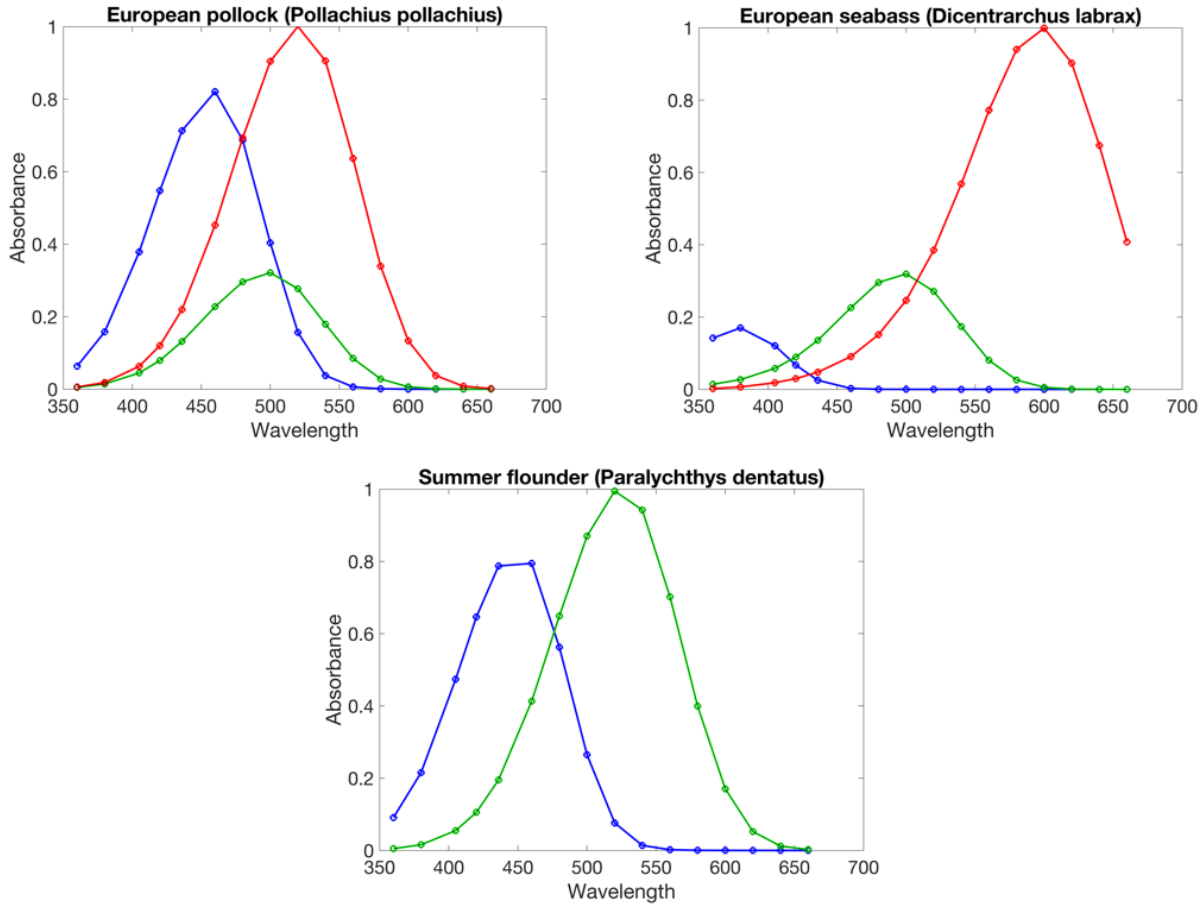


Figure 4.4: The spectral sensitivities for the three chosen fish predators. The spectral bands of the HSI camera are indicated by the circles. Top left: European pollock ($\lambda_{max} = 458, 498, 521nm$, Pigment weights $= 0.8214, 0.3214, 1$). Top right: European seabass ($\lambda_{max} = 380, 497, 598nm$, Pigment weights $= 0.17, 0.32, 1$). Bottom: Summer flounder ($\lambda_{max} = 449, 525$, Pigment weights $= 0.83, 1$)

Here, the units of divergence are in bits. A D_{KL} of zero indicates the two distributions are identical. Since the D_{KL} is not a symmetric value, the $D_{KL}(\Delta L_A||\Delta L_B)$ direction is chosen. In this assumption, the animal is treated as the "true" distribution and the background as the "approximate" distribution - this is consistent with the assumption that a visual predator searching for a cuttlefish compares every patch of background against a mental template of how a cuttlefish looks.

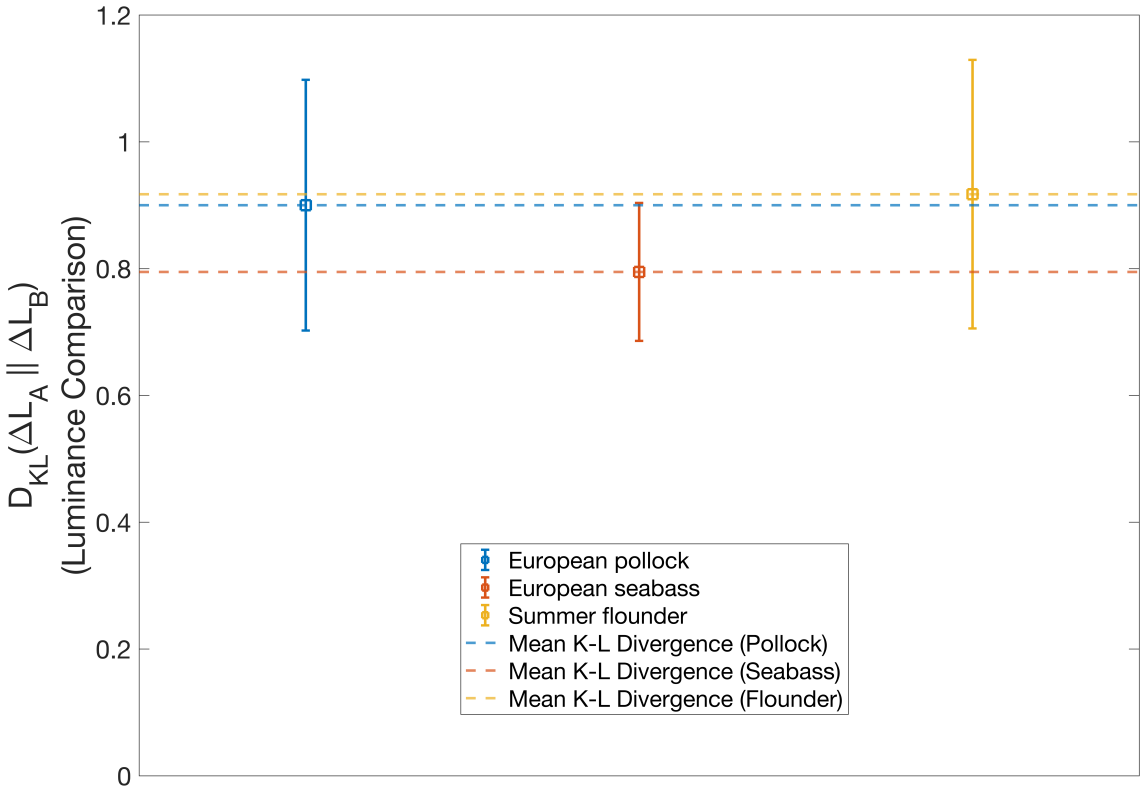


Figure 4.5: The distribution of $D_{KL}(\Delta L_A||\Delta L_B)$ for luminance for both predators (L_A is animal luminance, L_B is background luminance). The error bars give the standard error of the $D_{KL}(\Delta L_A||\Delta L_B)$. On the basis of this analysis, it cannot be claimed that discriminability via luminance is significantly different among the three predators, due to the overlap of error. However, mean values of $D_{KL}(\Delta L_A||\Delta L_B)$ for all three predators are higher than the mean values of chromatic discriminability (see Figure 4.7), suggesting that cuttlefish are more likely to be discriminable via contrast differences than color differences in all three predators.

In the luminance analysis, the distributions of $D_{KL}(\Delta L_A||\Delta L_B)$ are not significantly different for any of the predators (Figure 4.5). For all three of these predators, the diver-

gences in the luminance analysis are higher on average than in the color-match analysis (see Figure 4.7), suggesting that discrimination of the cuttlefish is more likely to occur via the identification of contrast differences than via the identification of color differences. This is consistent with the observation in Chiao, et al. [2] that the information in the luminance channel for potential predators is higher than in the chromatic channel.

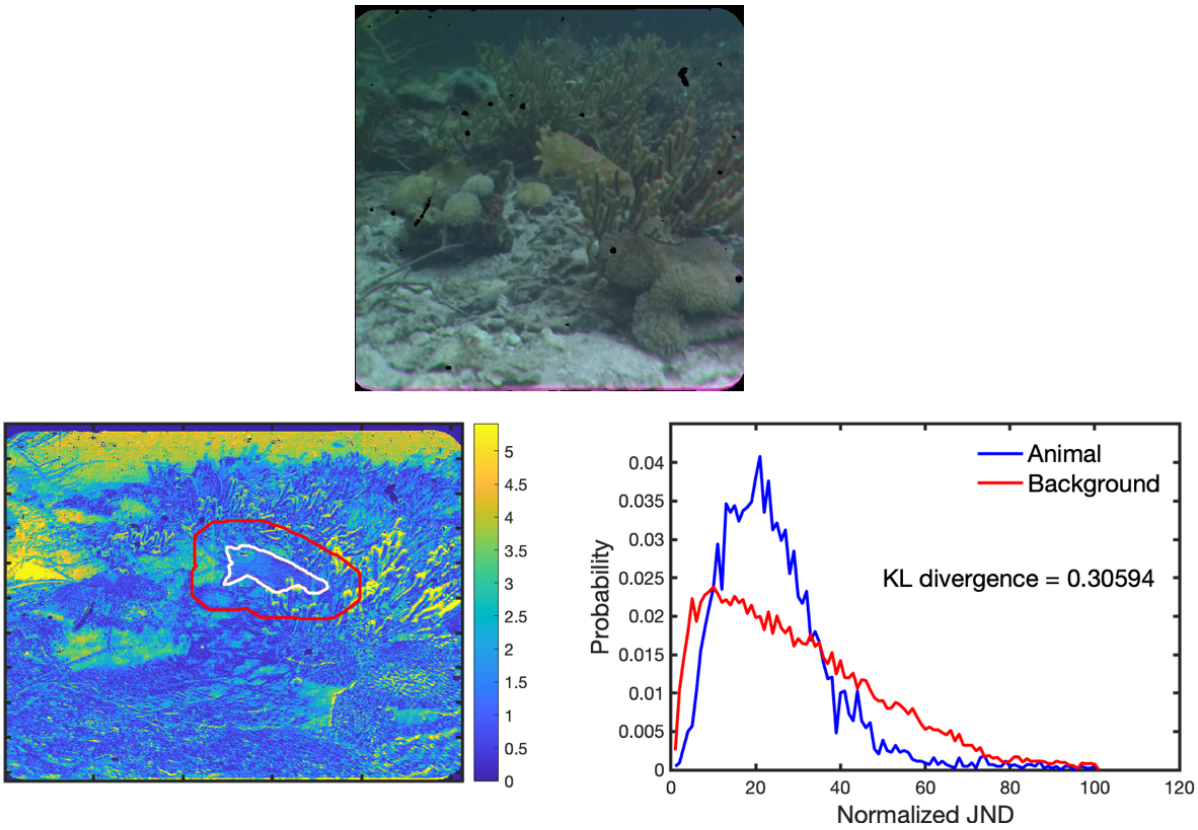


Figure 4.6: The JND map for a given image is shown from the perspective of the European pollock. The chosen background is the area between the red and white outlines, and the animal is outlined in white. A plot of the two JND distributions is also shown with a D_{KL} value.

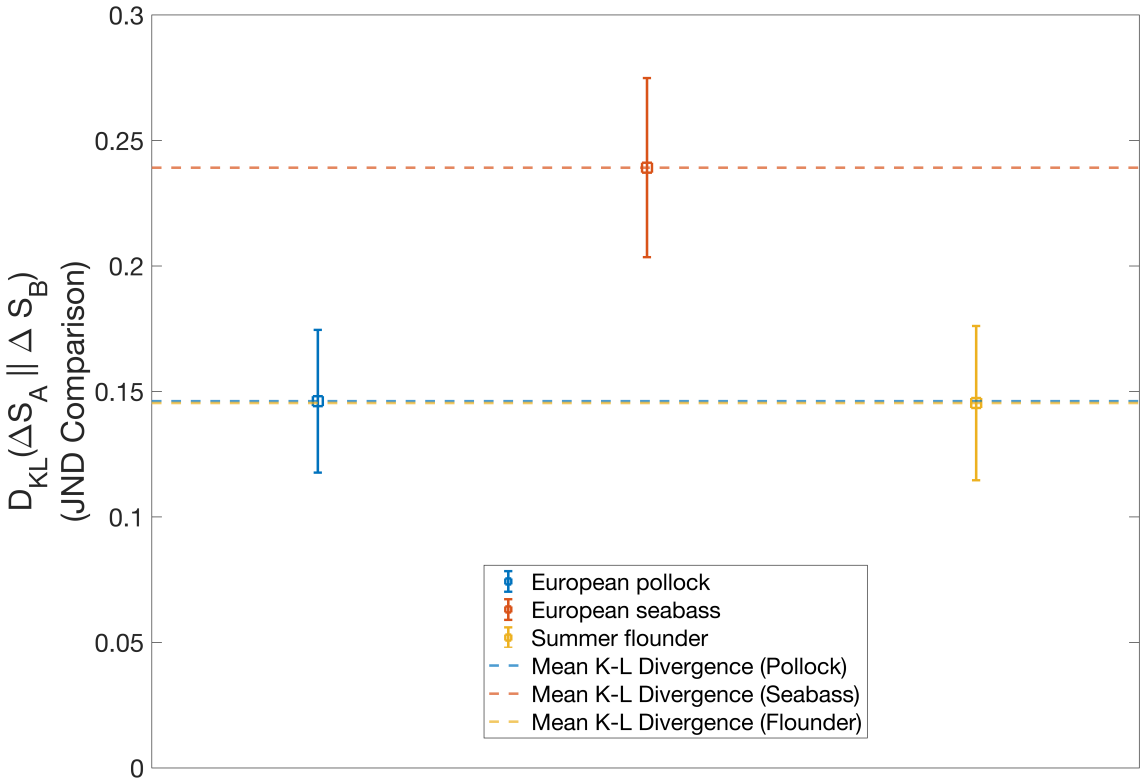


Figure 4.7: The distribution of $D_{KL}(\Delta S_A||\Delta S_B)$ for each fish predator, standard error on the mean plotted as error bars. On the basis of this analysis, the low values of $D_{KL}(\Delta S_A||\Delta S_B)$ indicate a good color match in the eyes of all three fish predators. There is a significant difference between chromatic discriminability in the seabass and the other two predators at a 5% significance level, suggesting that seabass are more likely to discriminate camouflage via color differences than the other two predators.

Now, more subtle features of cuttlefish camouflage, such as color-matching, can be investigated without considering luminance differences using the model developed by Vorobyev & Osorio [48]. In this strictly chromatic analysis, the distribution of just-noticeable-differences inside the animal (ΔS_A) can be calculated, and then compared to the distribution of just-noticeable-differences in the background (ΔS_B). An example JND map (from the perspective of the European pollock) for a given image is shown in Figure 4.6. The $D_{KL}(\Delta S_A||\Delta S_B)$ across the image dataset can then be calculated from the perspective of each predator, then plot the distributions along with standard errors. This plot is shown in Figure 4.7. The distributions of $D_{KL}(\Delta S_A||\Delta S_B)$ are restricted to lower bit values than in the luminance analysis, indicating a good color match in the eyes of all three predators and therefore difficulty of detection via chromatic differences. There is no significant difference in chromatic discriminability for the pollock and the flounder. The chromatic discriminability for the seabass is significantly different from that of the pollock (a two-sample T-test rejects the null hypothesis at a 5% significance level with a p-value of 0.0444) and from that of the flounder (a two-sample T-test rejects the null hypothesis at a 5% significance level with a p-value of 0.0496). This would suggest that seabass are more likely to discriminate camouflage via color differences than the other two predators.

In general, the lower divergence values in the chromatic analysis for these three predators might suggest that cephalopods have evolved a general strategy of color resemblance that works for the average fish visual system, regardless of the number of opsins or ecological niche. The significant difference between the seabass and the other two predators does, however, suggest that a visual system could be tuned to maximize chromatic discriminability. What might be responsible for the difference between the visual systems of the seabass and the other two predators? First, referring back to Figure 4.4, note that the peak total response of the European seabass is weighted further toward the red, compared to the peak total response of the flounder and the pollock, which actually have similar peak sensitivities in

their small and long-wavelength cones. Apart from the relative red-shift, if each pigment is weighted equally, the spread of peak cone sensitivities (standard deviation of 109.1) for the European seabass significantly outweighs the spread of peak sensitivities for the pollock (standard deviation of 31.88) and the flounder (standard deviation of 53.74). The direction of the shift and the relative spread could serve as starting points for a future project that tests the suggestion that visual systems could be tuned for maximum chromatic discriminability. A possible test would be scan the λ_{max} for two/three-pigment visual systems across a range of values corresponding to a range of shifts and spreads, calculating the $D_{KL}(\Delta S_A||\Delta S_B)$ across the image set for each new set of λ_{max} values. This optimization is beyond the scope of this thesis project, but is certainly a rich area for future investigation.

4.5 Discussion

In summary, the above results suggest that (1) cuttlefish have evolved to match color well in a generalist manner to be effective for different predator visual systems, and (2) cuttlefish are more detectable via contrast discrimination, and (3) it may be possible to tune the values of λ_{max} to maximize discriminability in the chromatic channel. First, it is important to place these results in the context of the landmark study on cuttlefish camouflage by Chiao, et al. [2], which also employed hyperspectral imaging. In this study, the authors [2] found that cuttlefish camouflaging on substrates in a lab environment matched the color and pattern of the background well in the perspective of hypothetical predator visual systems. The results of this chapter serve to support and reinforce this conclusion by showing that it holds under two key updates: (1) using images of cuttlefish camouflaging *in the wild* and (2) using *real* observed natural predators of the common cuttlefish. Earlier, I addressed the disconnect between the species in the pictures (*Sepia latimanus* and *Sepia apama*) and the species targeted in the wild by the selected fish predators (*Sepia officinalis*). The results above were collected with the assumption that these predators would still be well-equipped to hunt

other species of cuttlefish, but it would still be important to eliminate this limitation in future research by either collecting hyperspectral images of *Sepia officinalis* or by accelerating the effort of collecting visual data about predators of *Sepia apama* and *Sepia latimanus*.

In any case, combined with the Chiao et al. [2] study, these results further support the utility of hyperspectral in evaluating color vision in visual predators. The results of the seabass and flounder analyses open the door to further collections of actual cuttlefish predators whose visual systems are less well-characterized (especially *Serranus cabrilla* [5]). Furthermore, this analysis could easily be applied to other predator/prey pairings, whether above ground or underwater. For instance, a currently in-progress publication is investigating aposematic coloration in Pacific newts via hyperspectral imaging. It might also be instructive to repeat these analyses in the visual perspectives of strictly herbivorous or planktivorous fish that share ecological niches with cuttlefish. Additionally, it must be noted that studies characterizing the visual systems of natural predators are often separated by years or employ differing techniques (note the differences in methodology and publication times in the Horodysky, et al. [55] study and the Shand, et al. [56]) study. A possible solution to such discrepancies is to undertake a dedicated characterization study of a bank of cephalopod predators using a standard, modern procedure of photoreceptor characterization such as that employed in the Horodysky, et al. [55].

It would also be possible to test discriminability measures and models apart from the Vorobyev & Osorio [48] model used here. The possibility of live-animal experiments to test the HSI analysis is also exciting. The trainability of fish responding to visual stimuli presented on a computer monitor has some precedent in archerfish studies conducted by Ben-Simon, et al. [58]. Testing the responses of collected cuttlefish predators to black-and-white, full-color, and isoluminant target images of cuttlefish would allow for further investigation into the validity of hyperspectral imaging as a simulation method. Another rich area of future investigation is unlocked by the above results' implication that the values of λ_{max}

may be tuned to maximize discriminability in the chromatic channel. Generating a visual system optimized for chromatic discriminability might prove useful in testing the performance of artificial camouflage and cloaking - designing such technologies to be effective in this "theoretical best" visual system might optimize their performance in turn. Eventually, this optimized model could also have implications for technologies that seek to correct color-blindness or other color-related impairments. The potential future of hyperspectral imaging in the investigation of animal visual systems is extremely exciting, and only just beginning.

CHAPTER 5

CONCLUSION

This thesis aimed to address the challenge of quantifying camouflage by two means: the proposal of a new quantification metric, and the further examination of hyperspectral imaging to investigate color and luminance discriminability in predator visual spaces. Chapter 2 introduced the "texture distance metric", a measure founded on the loss function of a neural network-driven texture synthesis algorithm. It was also demonstrated how the metric can be used to generate "texture distance heatmaps" of scenes of cuttlefish camouflage, highlighting intuitive textural similarities and differences. In Chapter 3 tested the utility of this texture distance metric using a rich eye-tracking dataset collected on human subjects. It was hypothesized that any subject searching for a cephalopod will search the regions where the texture distance between cephalopod and background is low, and the hypothesis was confirmed by noting the proportion of fixations where the texture distance was below a given threshold. This analysis was refined by incorporating the duration of each fixation into the calculation, developing a metric called ρ_{img} to show that successful subjects will spend an even higher proportion of time fixating on low-texture distance values than failed subjects. Finally, in Chapter 4, chromatic/luminance discriminability analyses were performed through the perspectives of natural predators. To do so, hyperspectral images of cuttlefish taken under natural lighting conditions in the wild were used. The results suggest that cuttlefish have evolved to match color well in a generalist manner, and that they are more likely to be detected through luminance contrast rather than chromatic contrast. The results also suggest it may be possible to tune the peak sensitivities of a predator visual system to maximize discriminability in the chromatic channel.

The demonstration of the proposed texture distance metric not only contributes to the nascent research in quantifying camouflage, but also addresses the key question of predicting the visual fixations of predators in natural scenes. This exciting potential predictive ability

is only the first step, however - the texture distance metric can be further refined and tested with different image classifiers, more targeted layer selection, and larger and more diverse image sets. The analysis of hyperspectral images reinforces previous conclusions (Chiao, et al. [2]), while opening the door to exciting new possibilities for future research. The dataset can be strengthened with collection trips targeted toward measuring the spectral sensitivities of key cuttlefish predators (especially *Serranus cabrilla* [5]), improved forensic re-calibration techniques for the HSI camera, and even the observation of live-animal responses to contrast or color-tuned images of cuttlefish camouflage. These advances could in turn impact research into adaptable materials, medical imaging, and human vision. Such possibilities for future research directions are numerous, and it is exciting to imagine that this project constitutes yet another step toward a more complete understanding of the perception of cuttlefish camouflage.

APPENDIX A

FORENSIC RE-CALIBRATION OF THE HSI CAMERA

Due to a shift in the filter tray of the HSI camera during the collection trip, a significant percentage of data was corrupted. While the data in the analyses mentioned earlier were usable as is, there remains a number of images that need to be forensically re-calibrated. The progress and ongoing challenges in developing this re-calibration procedure are outlined in this appendix. A common calibration tool used during collections is a "color chart." These are palettes of different colors, each with known a reflectance spectrum - if the values of reflectance in an HSI image of a color chart match the known values of reflectance, the camera is well-calibrated. In the development of a forensic re-calibration procedure, HSI images of a Macbeth ColorChecker that were taken on certain days of the collection trip to Lembeh were used. The team at SurfaceOptics also provided a deconvolution file, which is supposed to adjust for cross-talk among the channels of the camera. However, the latest deconvolution file (referred to here as WH 4.5) does not adequately calibrate the camera, so this forensic recalibration procedure seeks to correct this deconvolution file. These two images of the Macbeth ColorChecker are "20190512124818.948_14ms.3d", here on referred to as "Indo_1", and "20190520093948.472_129ms.3d", here on referred to as "Indo_2". Pre-correction RGB representations of these two images are shown below for reference.

To do this, a function was written representing the mean square error between the mean white-balanced pixel values on the Macbeth Color Checker photographed in Indo_1 and the white-balanced true values of the color spectra. This function has 256 parameters, given by the 16 Macbeth colors and the 16 bands in each reflectance spectrum. Next, a gradient descent minimization is run on the WH 4.5 deconvolution file, constrained by this function. In other words, the program adjusts the WH 4.5 file such that the MSE between the values on the Macbeth chart in the image and the true values of the color spectra was minimized. Ideally, this should result in a deconvolution file that generates an RGB image

Figure A.1: Hyperspectral images of the Macbeth ColorChecker rendered as RGB images. These images are used to test a forensic recalibration procedure.

with the correct color spectra. Figure A.2 shows a heatmap representation of the WH 4.5 deconvolution file. Figure A.3 shows the performance of WH 4.5 on the Macbeth spectra alongside the true values, and Figure A.4 shows an RGB representation of Indo_1 generated using the WH 4.5 file.

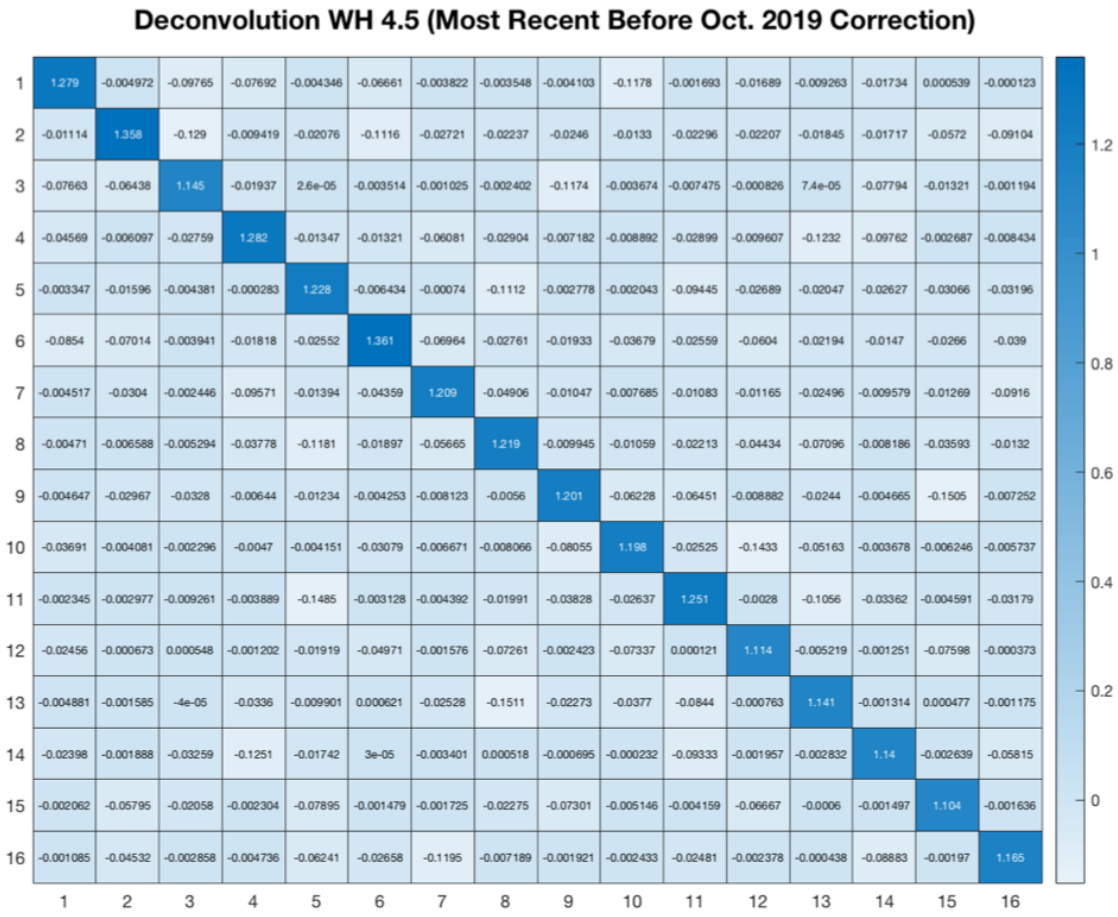


Figure A.2: A heatmap representation of the initial, pre-correction deconvolution file (WH 4.5). Since the goal of the deconvolution file is to correct cross-talk among the 16 color channels of the HSI camera, its dimensions are 16x16.

As observable from the Macbeth spectra performance, the measured spectra are extremely noisy, and by and large do not align with the true values. Following the optimization using only the spectral data from Indo_1, the heatmap of the corrected deconvolution matrix also changes, indicating much greater cross-talk, particularly in Band 5, corresponding to 436 nm. The heatmap for the corrected deconvolution matrix, is shown in Figure A.5. The performance also improves significantly, as shown in Figure A.6. An RGB representation of Indo_1 generated using the optimized deconvolution file is shown in Figure A.7.

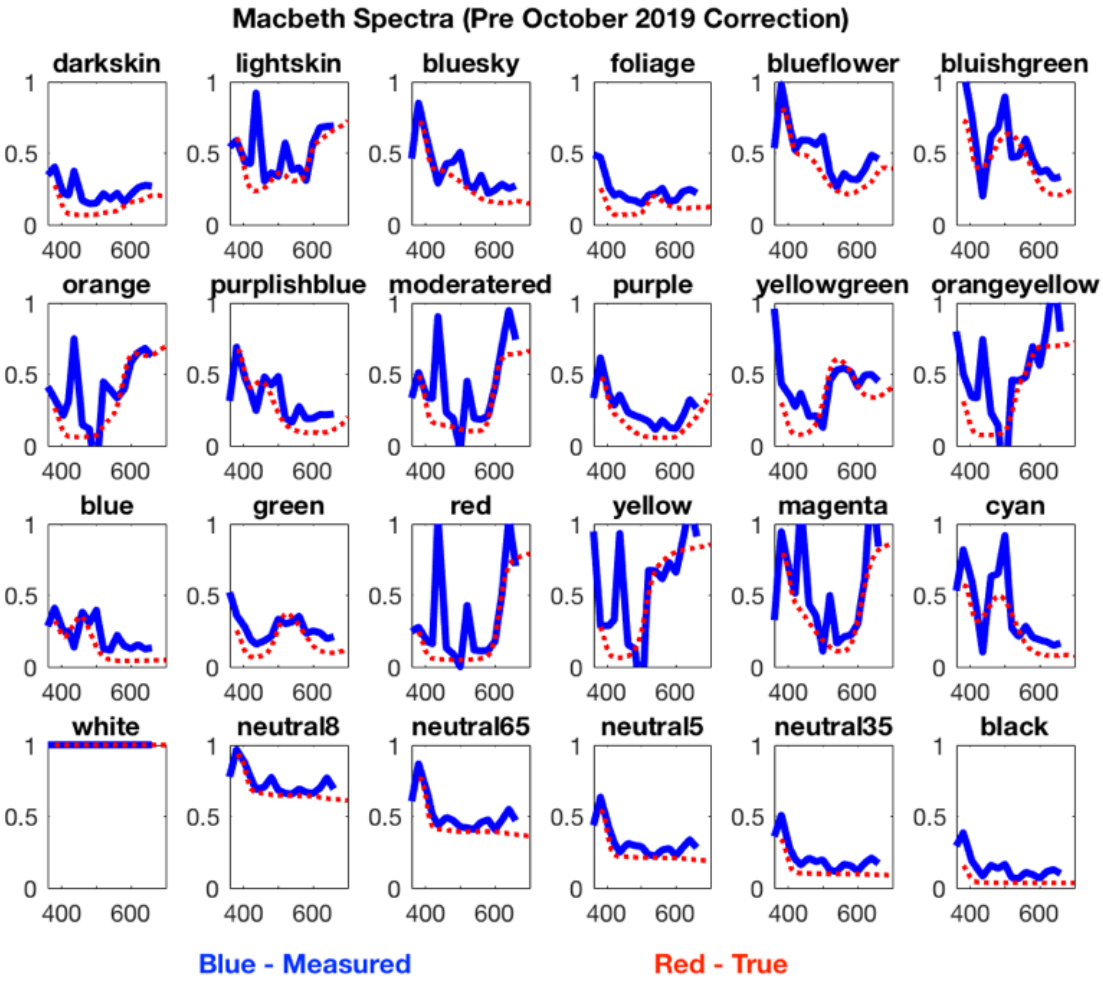


Figure A.3: Using the pre-correction deconvolution file on Indo_1 results in poor performance: the measured reflectance spectra do not match the true reflectance spectra.

Figure A.4: The RGB representation of Indo_1 generated using the intial, pre-correction deconvolution file.

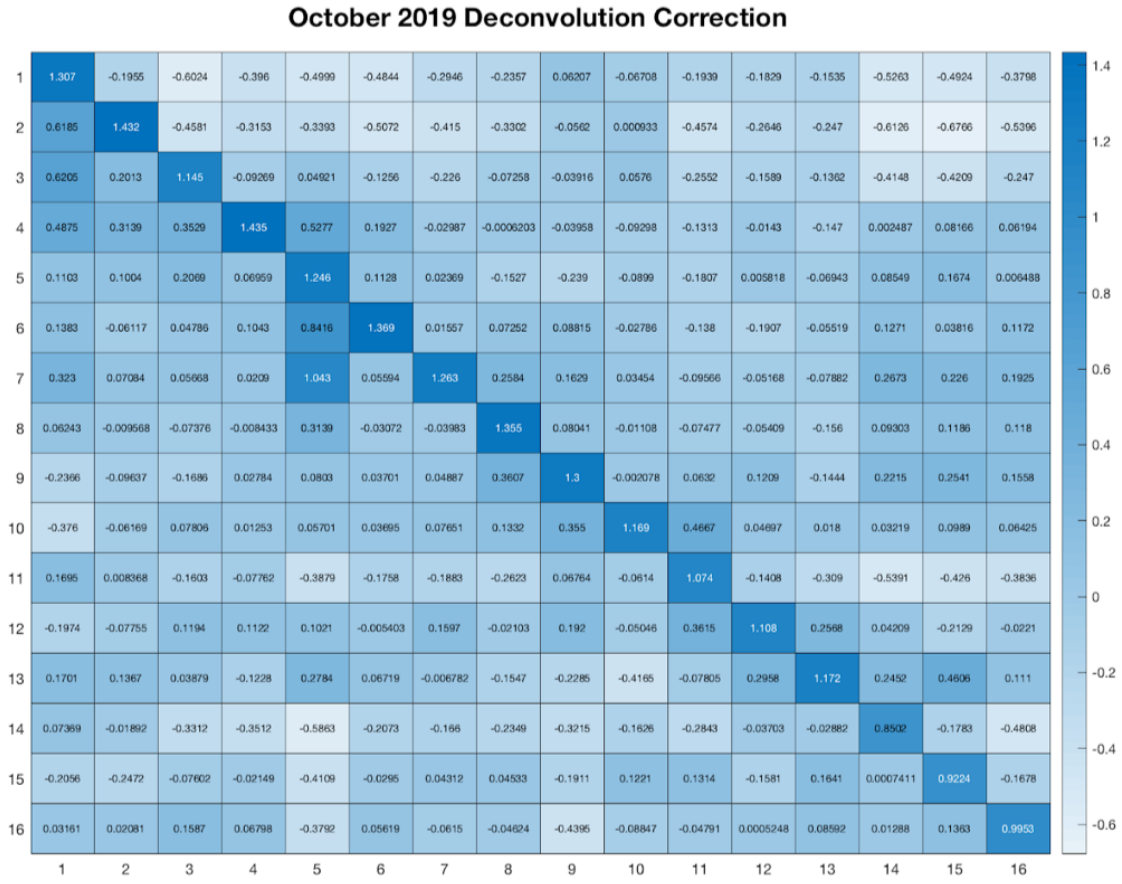


Figure A.5: A heatmap representation of the deconvolution file after correction with Indo_1 as a constraining image.

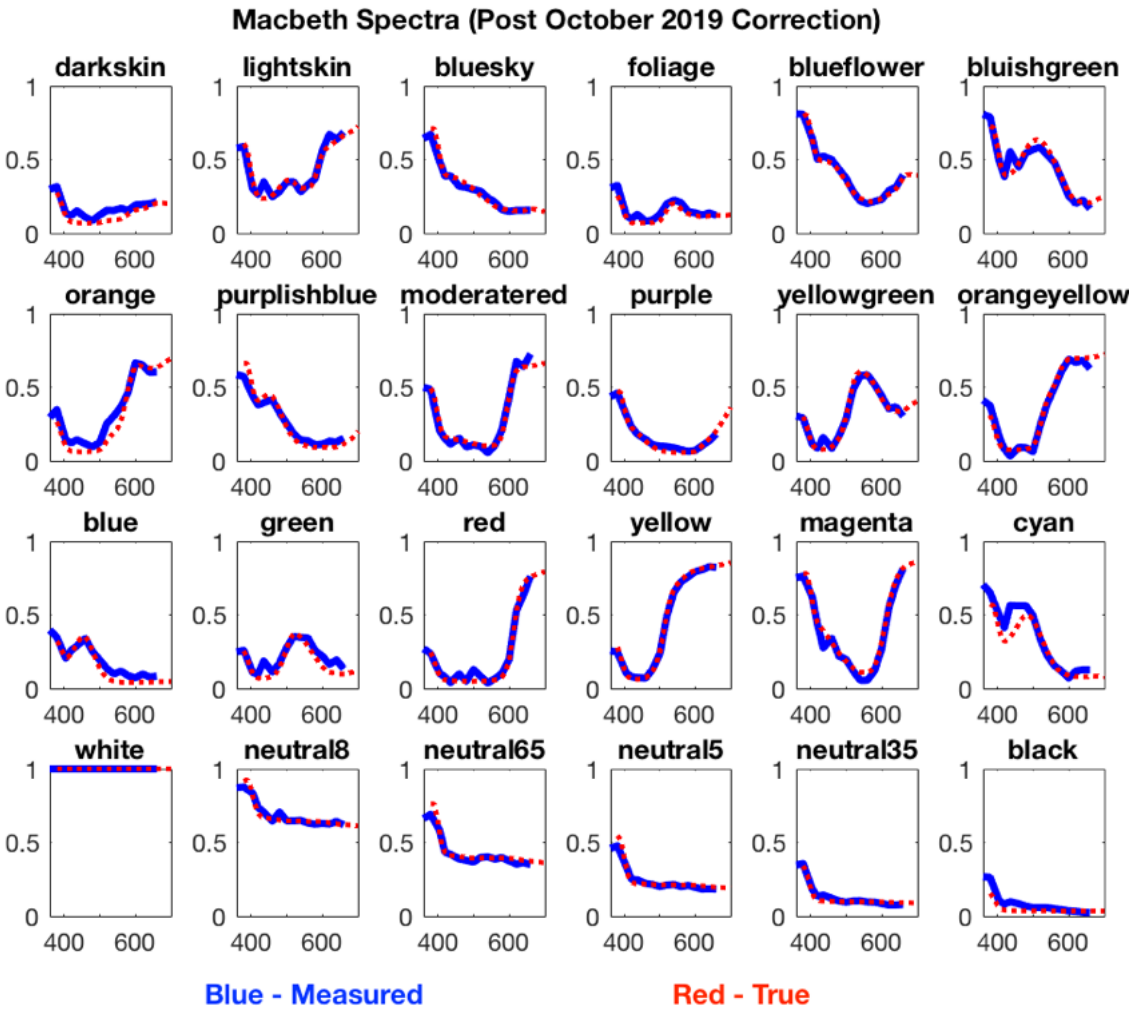


Figure A.6: Using the Indo_1-constrained corrected deconvolution file on Indo_1 results in excellent performance: the measured reflectance spectra matches the true reflectance spectra.

Figure A.7: An RGB representation of Indo_1 generated using the Indo_1 constrained corrected deconvolution file.

It must, however, be noted that this optimization only took into account spectral data from a single image, namely Indo_1. As such, it is very likely that this optimization is at least somewhat image specific. Indeed, when the corrected deconvolution matrix is applied to a different Macbeth chart image, "20190520093948.472_129ms.3d" (Indo_2), it is found that the stellar performance on the Macbeth spectra does not quite hold up, as shown in Figure A.8. An RGB representation of Indo_2 using the corrected deconvolution file is shown in Figure A.9.

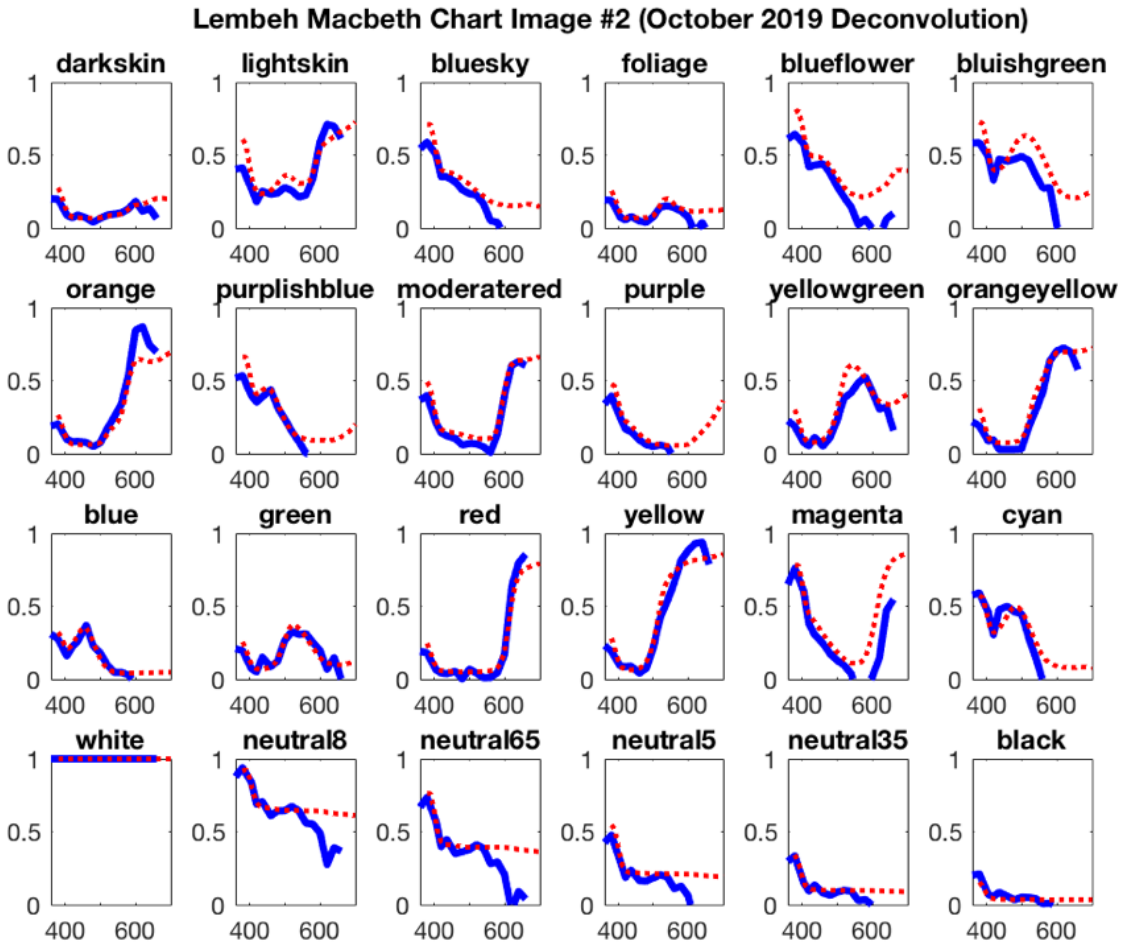


Figure A.8: Using the Indo_1-constrained corrected deconvolution file on Indo_2 results in poor performance: the measured reflectance spectra do not match the true reflectance spectra

Figure A.9: An RGB representation of Indo_2 generated using the Indo_1 constrained corrected deconvolution file.

While not nearly as poor as the pre-correction results, it cannot be claimed that the measured data aligns with the true data in the case of the second image. To remedy this issue, a version of the optimization program that accepts multiple images as input was written, such that spectral data from a wide variety of images could be taken into account. To start, the optimization was run with both Indo_1 and Indo_2 as constraints. The heatmap of the resulting deconvolution matrix is shown in Figure A.10. The performance of this multi-image-constraint heatmap on the Macbeth spectra for Indo_1 and Indo_2 is shown in Figures A.11 and A.12., and the corrected RGB representations of the two images is shown in Figure A.13.

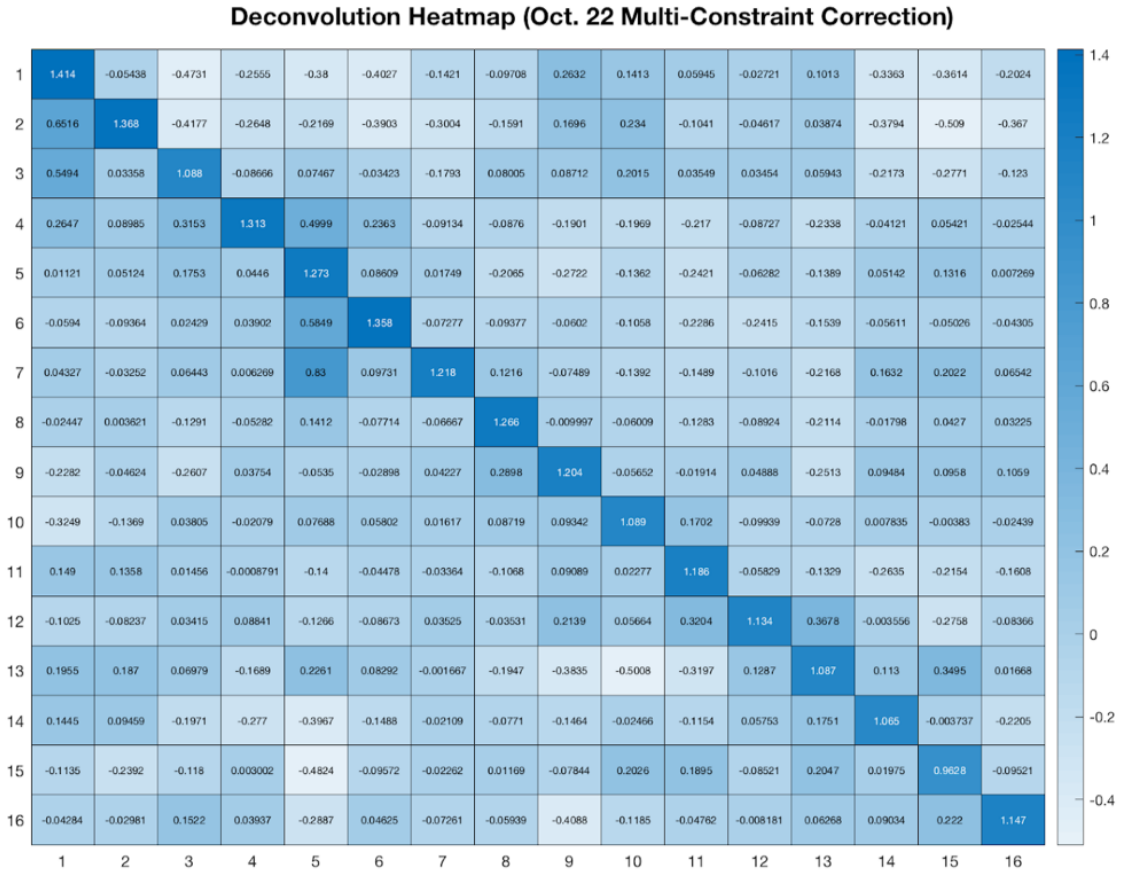


Figure A.10: A heatmap representation of the deconvolution file after correction with Indo_1 and Indo_2 as constraining images.

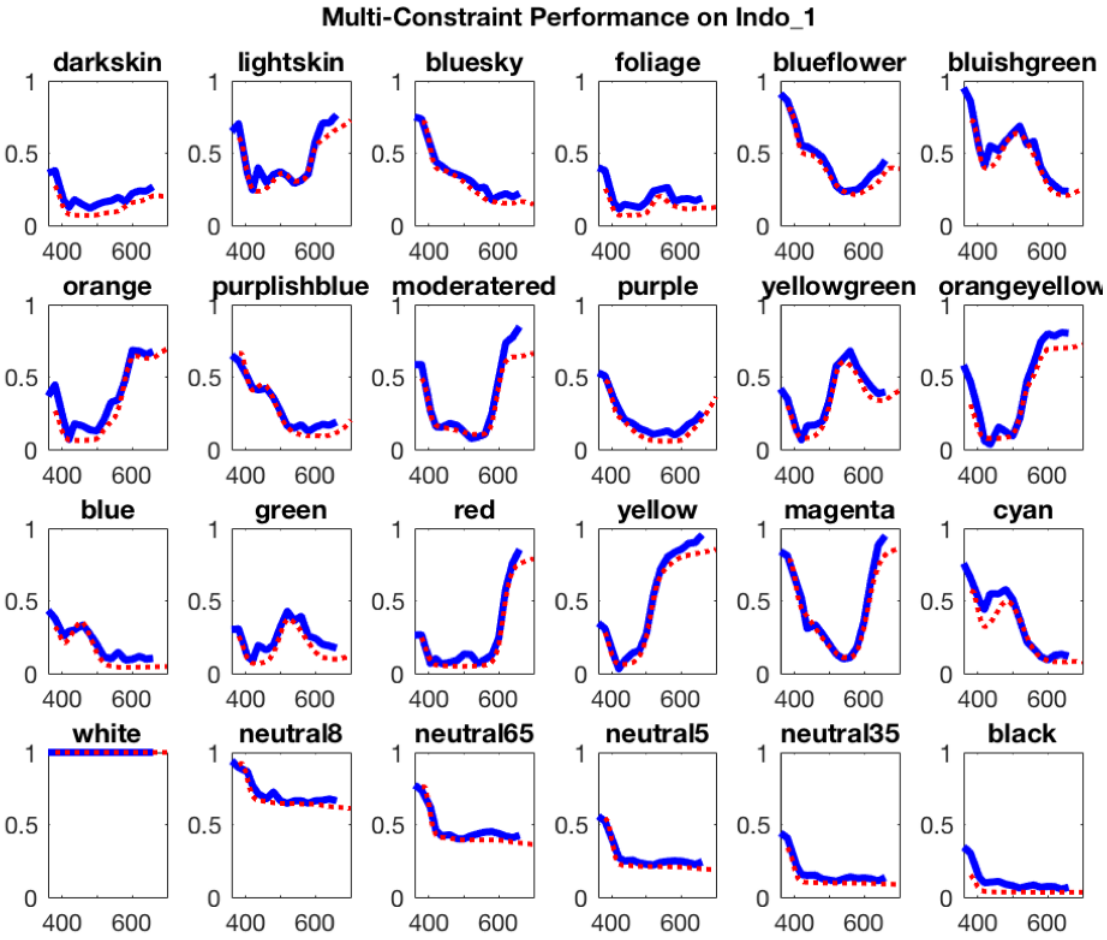


Figure A.11: Using the multi-image-constrained corrected deconvolution file on Indo_1 results in good performance: the measured reflectance spectra matches the true reflectance spectra

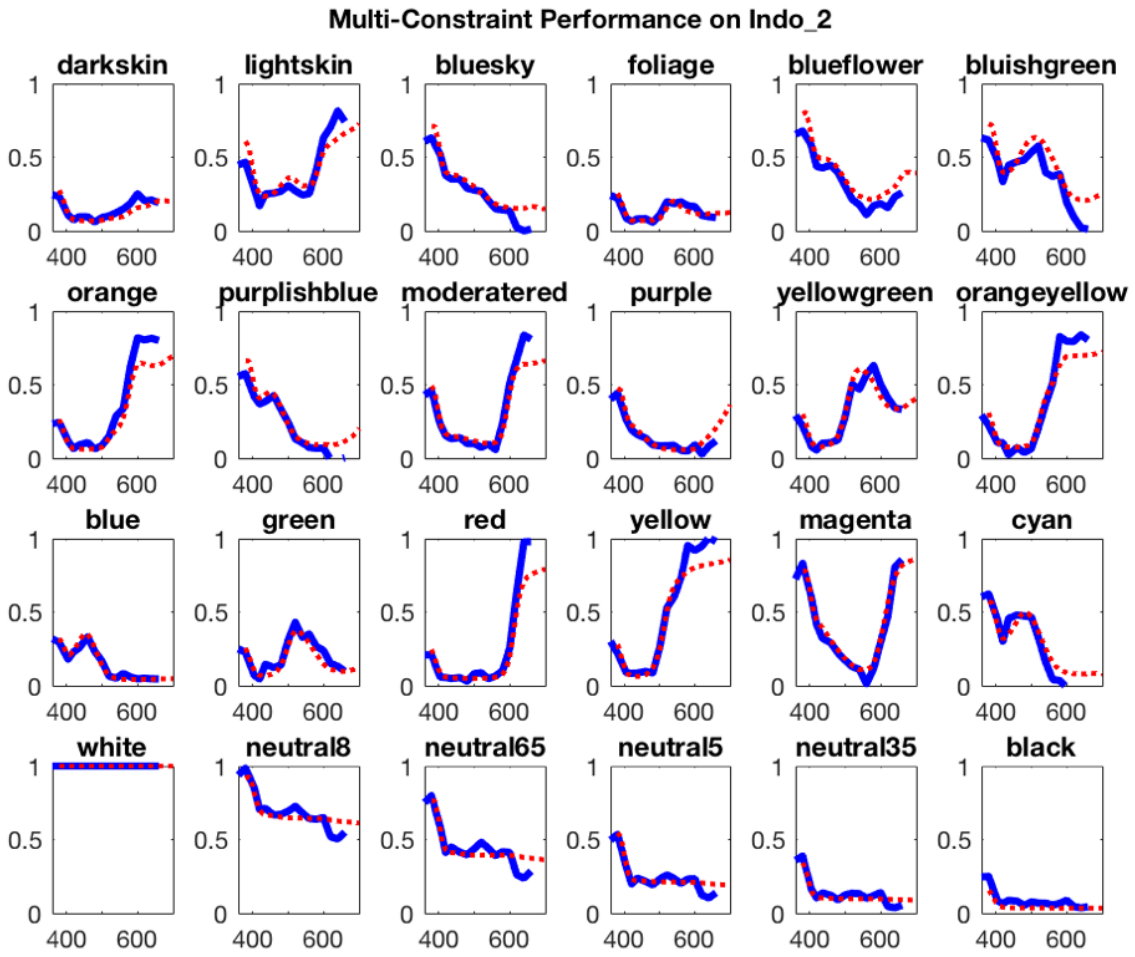


Figure A.12: Using the multi-image-constrained corrected deconvolution file on Indo_2 results in good performance: the measured reflectance spectra matches the true reflectance spectra

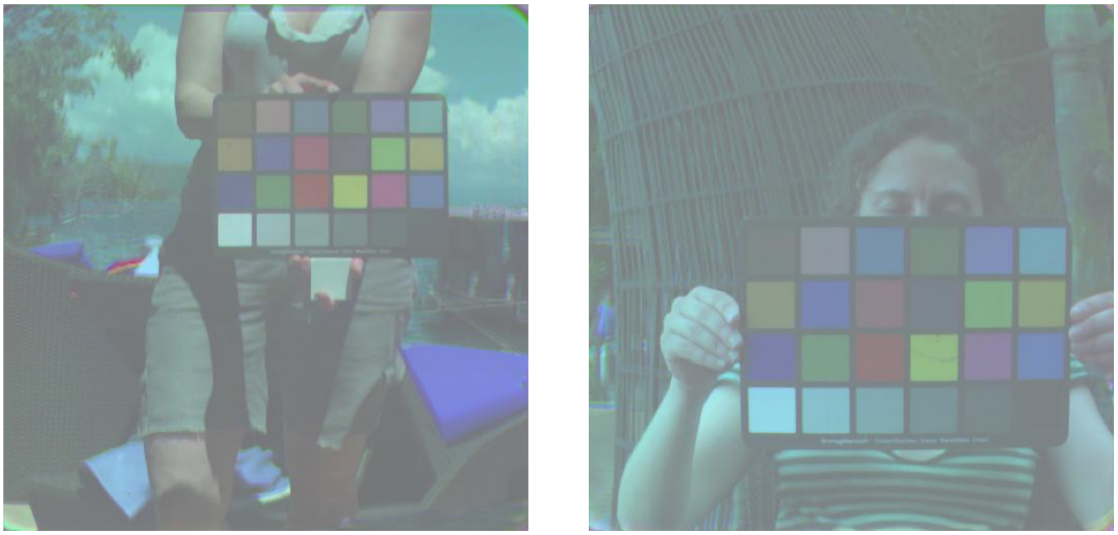


Figure A.13: RGB representations of Indo_1 and Indo_2 generated using the multi-image-constrained corrected deconvolution file.

This averaging approach does in fact significantly improve the quality of Macbeth spectra performance on Indo_2 compared to the single-image constraint. However, it must be noted that the Indo_2 performance is still not perfect and, in the process of averaging, the quality of Indo_1 performance has been compromised slightly. The Indo_1 and Indo_2 spectra also do not have the exact same tendencies, though they generally preserve the correct shape of the true values. The RGB representations appear less washed out in the averaging approach than in the single-image constraint. With all those limitations noted, the averaging approach has yielded results which far exceed the quality of those obtained using the pre-correction WH 4.5 file. Furthermore, this averaging approach has so far only been carried out on a sample size of two images. It is suspected that with even more sample images and lighting conditions, the Macbeth spectra performance should be correspondingly more faithful to the true values and more reproducible across different images. However, a few more tests and checks of the procedure are necessary.

A heatmap showing the differences between the pre-correction and the Indo_1-corrected version is shown in Figure A.14. Based on the poor performance of the pre-correction file

on the Macbeth Spectra across a wide range of wavelengths, the optimization process would have to correct intensities for a wide range of wavelengths, so it makes sense that the matrix of differences would heat up across the spectrum as well. This is also in line with the observed shift of the filter tray.

An image from 2018 (20180628072051.871_4ms.3d), from before the camera was damaged, was used as a control. For this image, the WH 4.5 file performs decently on the Macbeth Spectra test, with the strong exception of the bands below 450 nm, which the SurfaceOptics technician with whom the team corresponded changed by hand. This is shown in Figure A.15. If the optimization process is doing what it should, if 20180628072051.871_4ms.3d is now used as the constraining image, the corrected deconvolution file should primarily deviate from identity only in the first few bands, and the differences matrix should similarly heat up primarily in the first few bands. It seems that these results confirm expectations, as shown in Figure A.16. The optimization method appears to not only correct dramatically incorrect deconvolution files, but it also seems to be able to hone in on specific cross-talk and not over-correct already decent deconvolution matrices.

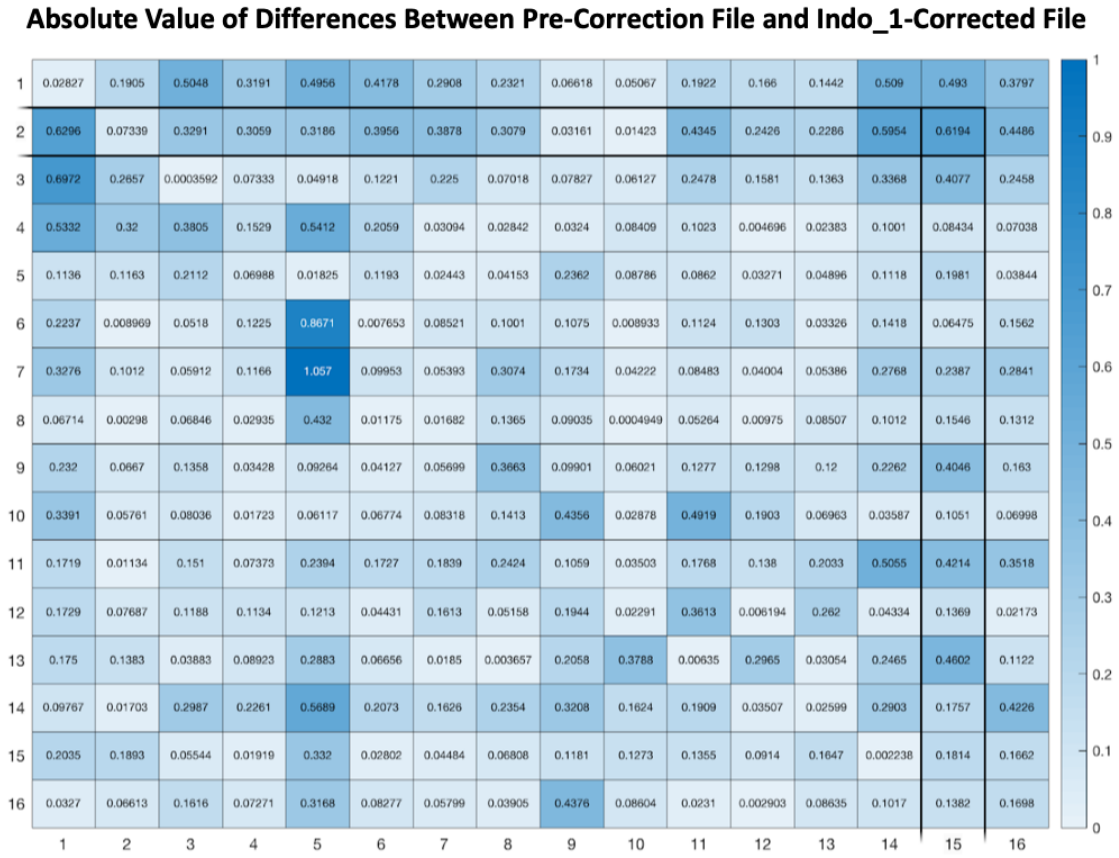


Figure A.14: A heatmap representation of the absolute value of differences between the pre-correction deconvolution file and the Indo_1-constrained corrected deconvolution file. Note that correction affects a wide range of spectral bands in this case.

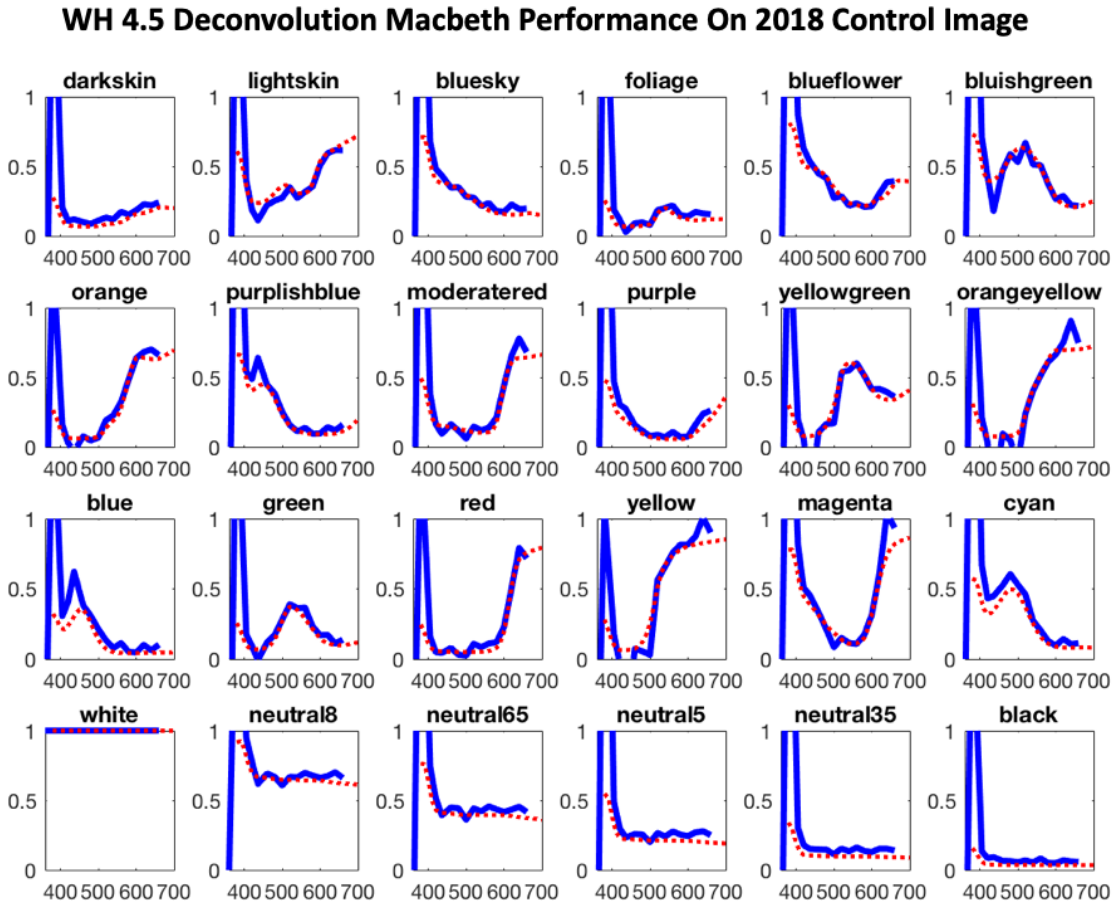


Figure A.15: Using the pre-correction deconvolution file on the 2018 control image results in good performance with the exception of the early bands, which were changed by hand by the SurfaceOptics technician.

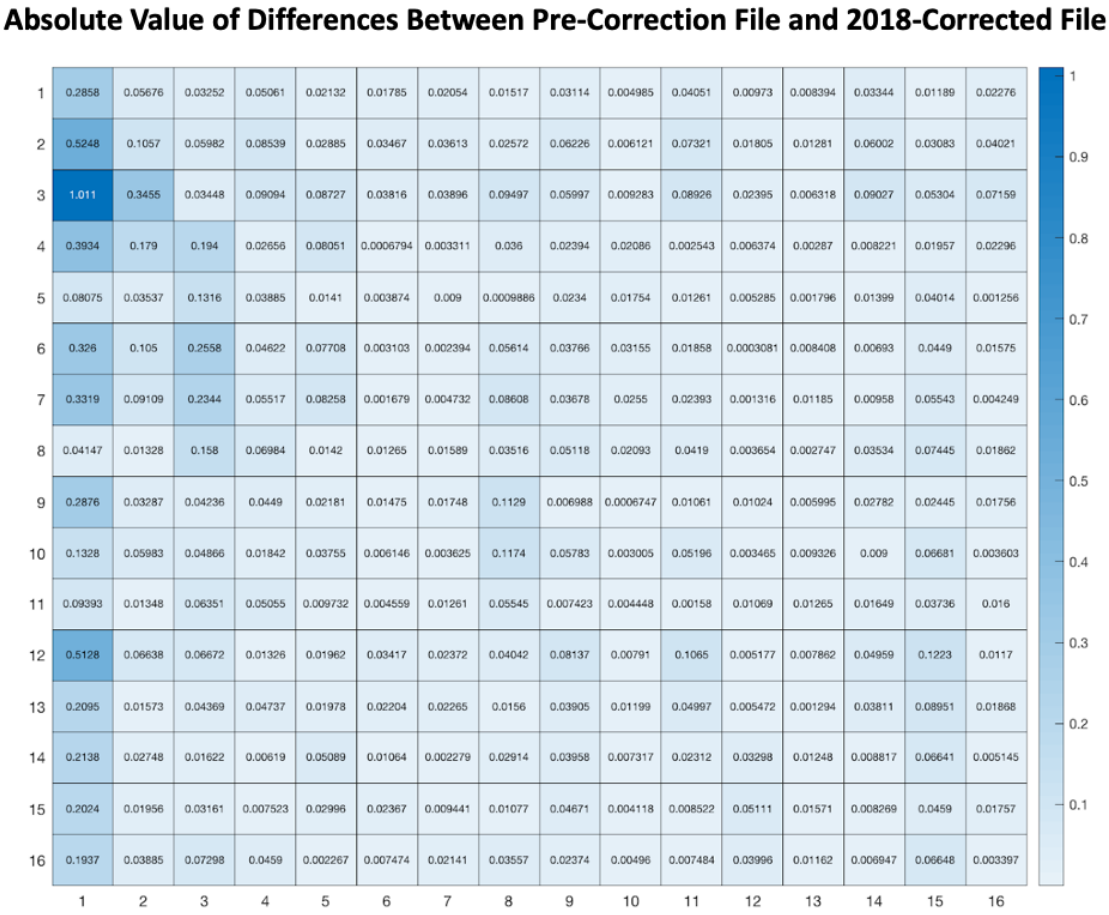


Figure A.16: A heatmap representation of the absolute value of differences between the pre-correction deconvolution file and the 2018 control-constrained corrected deconvolution file. Note that correction affects a much narrower range of spectral bands in this case, as expected.

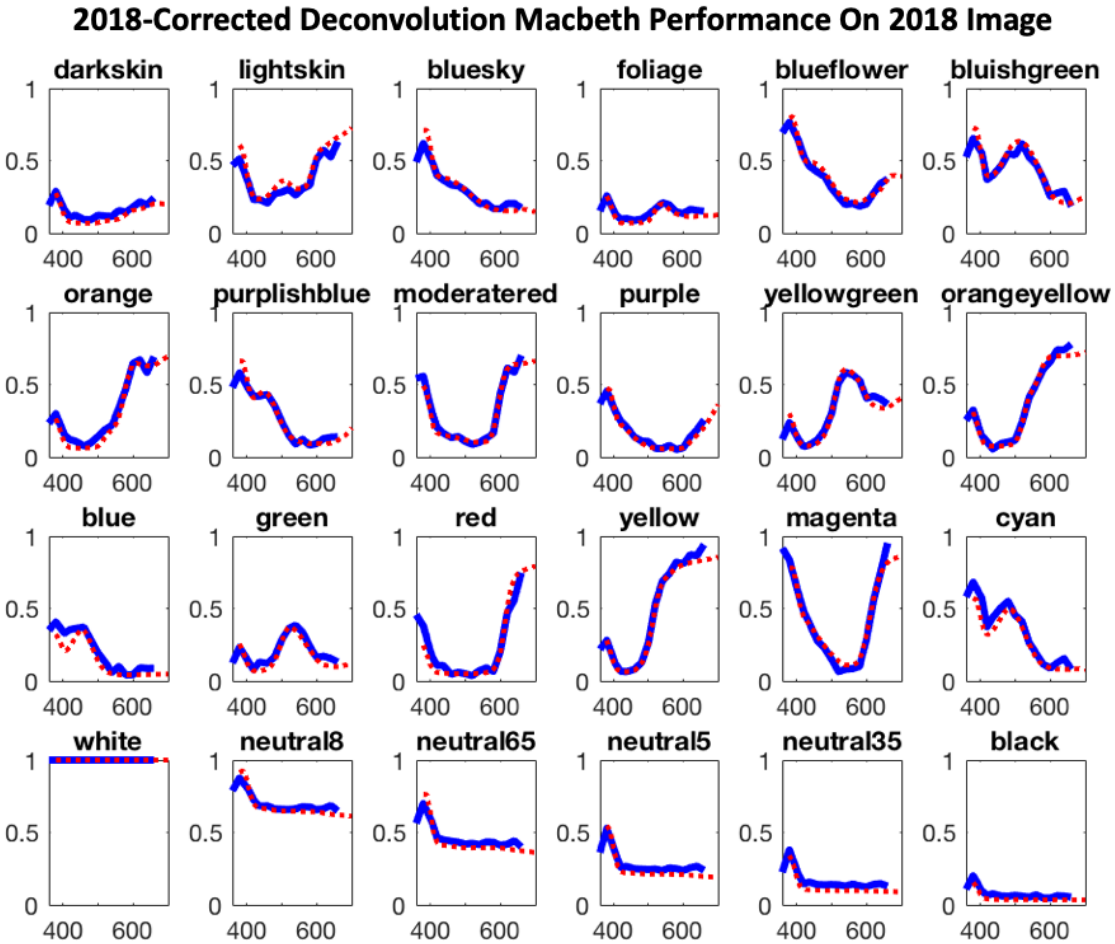


Figure A.17: Using the 2018 control-constrained deconvolution file on the 2018 control image results in excellent performance.

While this method would be very useful for future trips, it requires daily images of the Macbeth Color Checker, since it is now understood that the filter array can shift over time. This presents a challenge for past trips, and particularly for the Raja Ampat trip, in which an image of the Macbeth Color Checker was only taken on some days. Images of a Spectralon white standard were, however, taken underwater every day. To this end, a collaboration with Dr. Derya Akkaynak aims to rework this method such that it optimizes by using the Spectralon as a constraint. In the meantime, to investigate whether optimization can take place using only whites/neutrals as a constraint at all, the method is run on two specific Macbeth chart images taken during the Lembeh trip. These images are shown in Figure A.18.

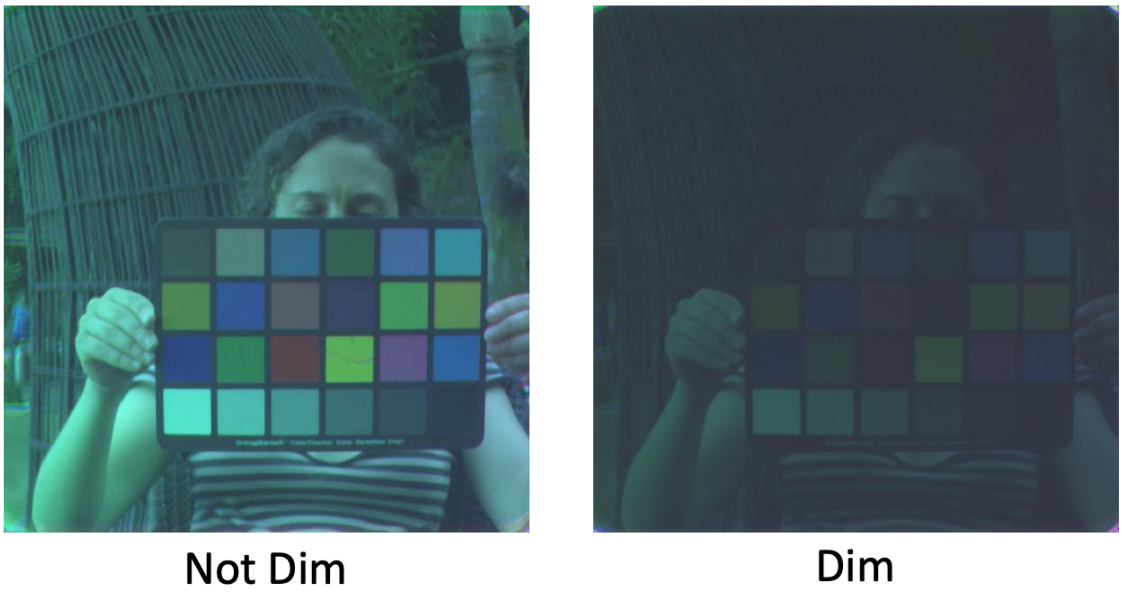


Figure A.18: Images used to investigate whites/neutrals-only optimization. These images are essentially the same photograph with different exposure times, and are therefore ideal for this experiment.

These two images are essentially the same photograph taken with two different exposure times, the same set of masks can be used for both and all variables between the two can be eliminated except for brightness. It was seen earlier that if optimization used data from only one image, it faithfully reproduced the true Macbeth spectra for that image, but only

somewhat or imperfectly for an image on which it was not trained. In this case, with the Dim/Not Dim images, if training is based on one of these images, it will similarly reproduce the true Macbeth spectra for that image. However, now it can be assumed that the differences between the measured and true spectra for the image it was not trained on will be solely due to the difference in brightness between the two images. If those imperfections can then be corrected by cotraining with just the whites/neutrals in the other image, or by applying some multiplicative white-balancing factor, that would suggest that Raja Ampat data can indeed be recovered using just the Spectralon white chart. First, two deconvolution files are created, one trained on the Dim image, and one trained on the Not Dim image, using the WH 4.5 Deconvolution as a starting point.

Notice that for both the Dim and Not Dim images, the deconvolution file trained only on the Not Dim image performs better than the charts generated using the old WH 4.5 deconvolution file (Figure A.19). However, note that the chart for the Dim image obtained with the Not Dim deconvolution (Figure A.19, bottom right) is still far from perfect. In particular, this chart appears to show a number of noisy peaks and diverges from the target at many points. As detailed above, it is reasonable to assume that this difference is due to the difference in exposure between the two images. The question then becomes, how can only the data from the white/neutral patches of these images be used to change the Not Dim deconvolution file such that a deconvolution file that works for the Dim image is obtained? If this is possible, then it may be possible to use the Spectralon white standard data from each day of the Raja Ampat trip, along with the Macbeth data from certain days, to recover that data.

The first approach was to train on all the colors of the Not Dim image, then use the resulting deconvolution file as the starting point for a new optimization that would only use the neutrals (white, neutral8, neutral65, neutral5, neutral35) from the Dim image. The results of this process are shown in Figure A.20. Unfortunately, with this approach, although

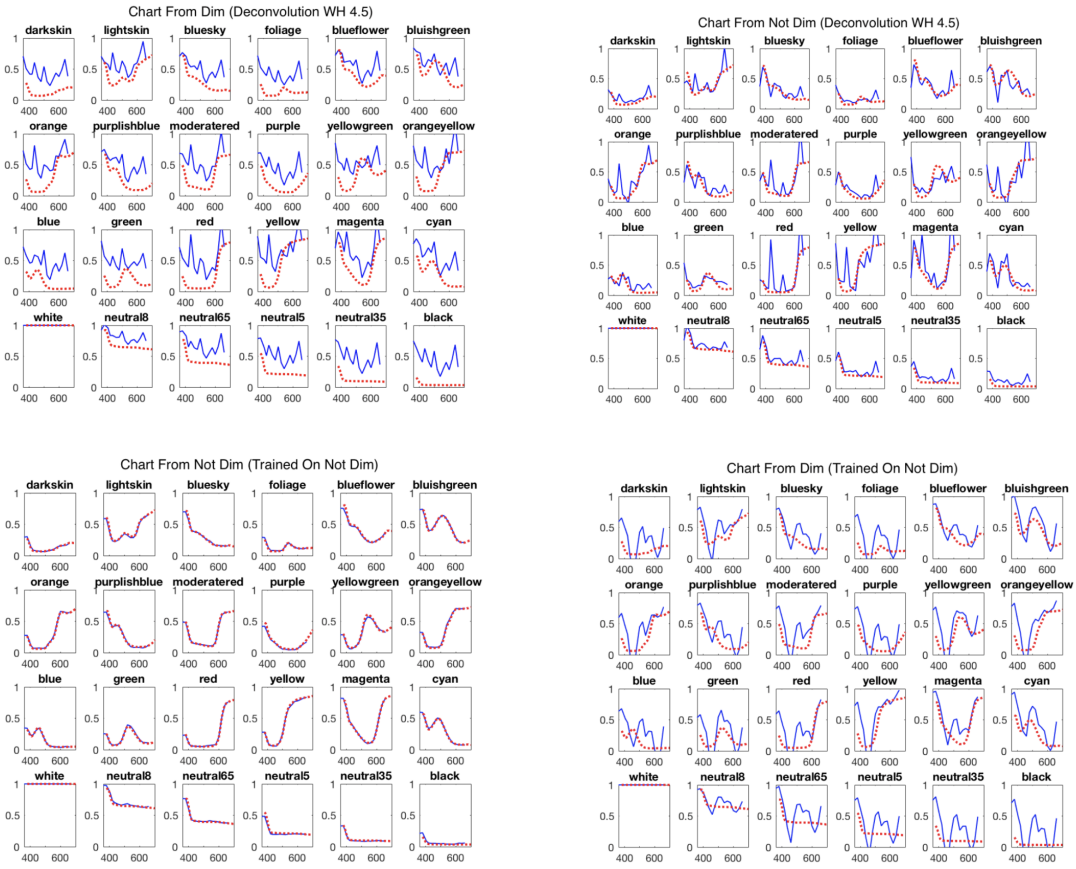


Figure A.19: Using the Not-Dim-constrained correction on the Not-Dim image results in excellent performance compared to using the pre-correction deconvolution file. Using the Not-Dim-constrained correction on the Dim image results in slightly better performance compared to using the pre-correction deconvolution file.

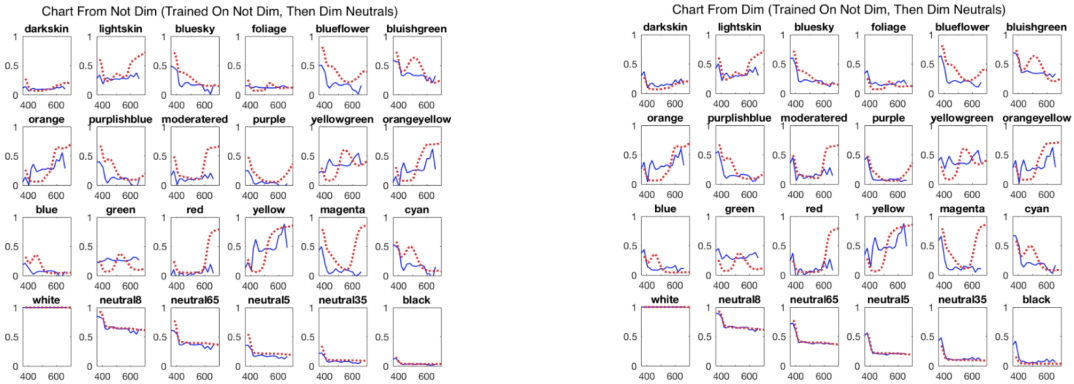


Figure A.20: Performance with two-step optimization (train on Not Dim, then Dim neutrals). Vertical shift appears to be corrected, but shapes of non-neutral curves are not preserved.

the vertical shift seems to vanish, the shapes of most of the non-neutral colors also seems to vanish. The next approach was to co-train, using WH 4.5 as a starting point. In other words, I would use all the colors of the Not Dim image, and just the neutrals of the Dim image as constraints in a single optimization rather than two back-to-back optimizations with separate constraints. The results of this co-training approach are shown below.

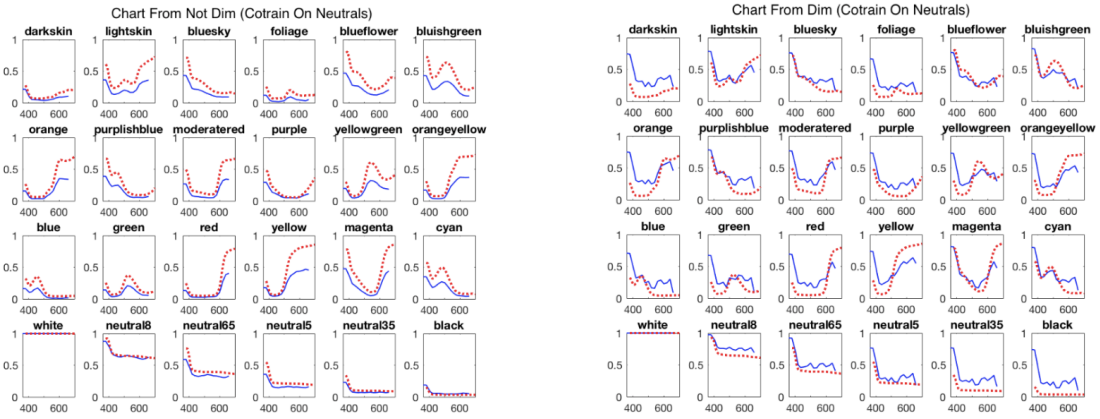


Figure A.21: Performance with co-training optimization (train on Not Dim and Dim neutrals simultaneously). Vertical shift appears to be corrected, but shapes of non-neutral curves are not preserved.

Similar to the previous approach, co-training also seems to destroy the shapes of the Dim colors, and introduces noisy peaks in the neutrals. Next, the difference between the raw white spectra of the Dim and Not Dim images is investigated. This is plotted in Figure A.22. The Y-axis here appears suspiciously small, and indeed, when the *true* raw white spectrum of the Macbeth ColorChecker is plotted on the same plot, it dwarfs the dim/not dim white spectra, as shown in Figure A.23. The same process is then repeated for the Darkskin color patch to confirm this, and the same result is seen in Figure A.24.

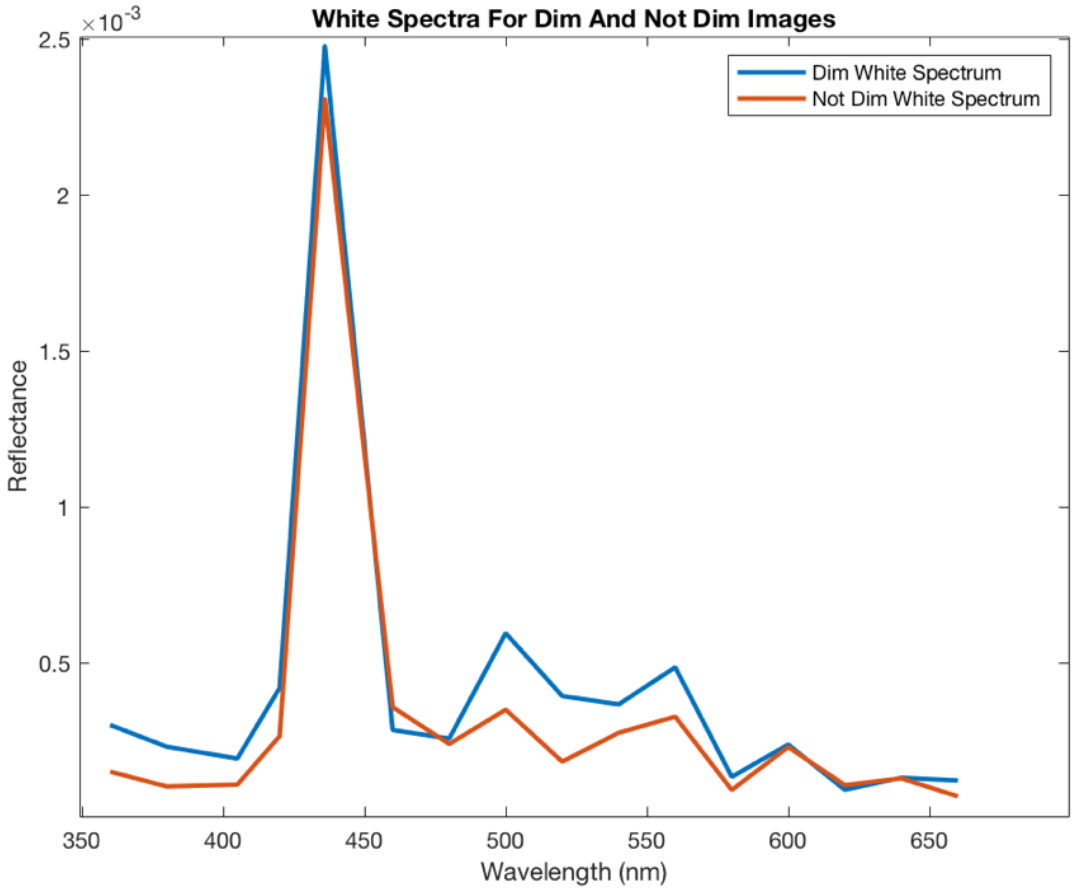


Figure A.22: Raw white spectra of Dim and Not Dim images.

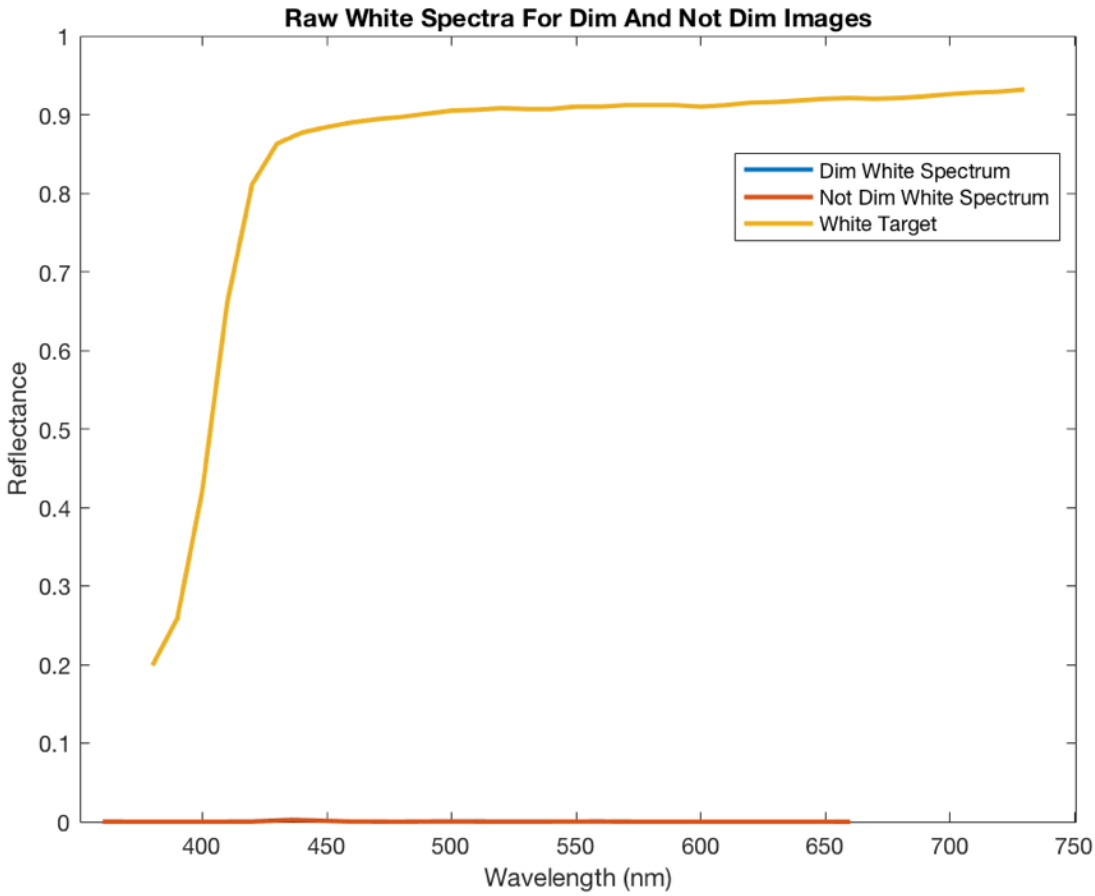


Figure A.23: Raw white spectrum of Macbeth ColorChecker is orders of magnitude higher than raw white spectra of Dim/NotDim images.

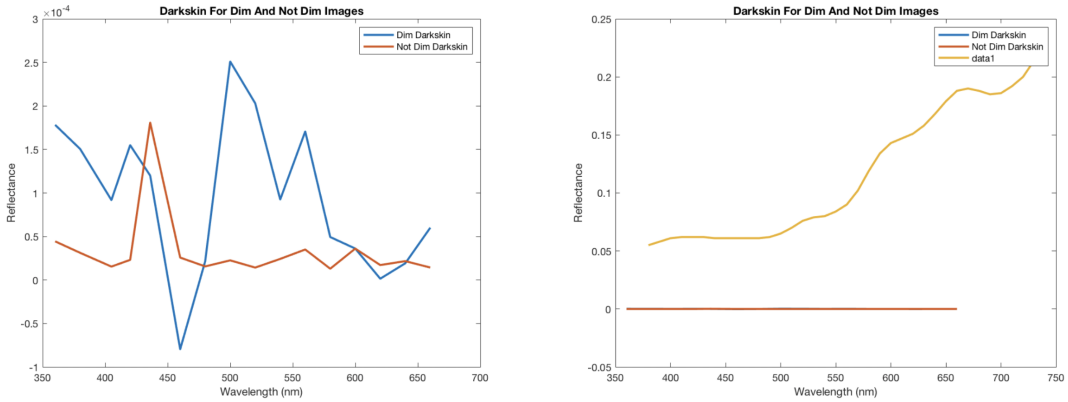


Figure A.24: Raw darkskin spectrum of Macbeth ColorChecker is also orders of magnitude higher than the raw darkskin spectrum of the Dim/Not Dim images. This suggests optimization should first take place without white-balancing.

This investigation led to the realization that the differences between the white-balanced spectra shown in Figures A.23 and A.24 may be due to the fact that the raw spectra outputted by the corrected deconvolution file have a significantly small magnitude in comparison to the true raw spectra. In other words, it is suspected that even a small shift in the raw spectrum at that magnitude would translate to a large shift after white-balancing, in a way that could inform the differences seen between the dim and non-dim images. The initial program minimizes the mean-squared error between the measured white-balanced spectra and the white-balanced truth, since these were most often used in the analyses. However, it is now clear that this is no guarantee that the raw spectra will be close to the raw target, and it is difficult to predict how the differences between raw spectra on the order of magnitude 10^{-4} will translate after white-balancing with a raw white spectra on the same order of magnitude. Therefore, I propose that the correct path forward is to write a version of the program that optimizes the mean-squared error between the measured raw spectra and the target raw spectra, and see how this might affect the differences between the dim and not-dim charts and the Indo_1 and Indo_2 analyses.

REFERENCES

[1] Leon Gatys, Alexander S Ecker, and Matthias Bethge. Texture synthesis using convolutional neural networks. *Advances in neural information processing systems*, 28, 2015.

[2] Chuan-Chin Chiao, J Kenneth Wickiser, Justine J Allen, Brock Genter, and Roger T Hanlon. Hyperspectral imaging of cuttlefish camouflage indicates good color match in the eyes of fish predators. *Proceedings of the National Academy of Sciences*, 108(22): 9148–9153, 2011.

[3] Martin Stevens and Sami Merilaita. *Animal camouflage: mechanisms and function*. Cambridge University Press, 2011.

[4] Roger T Hanlon and John B Messenger. Adaptive coloration in young cuttlefish (sepia officinalis l.): the morphology and development of body patterns and their relation to behaviour. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 320(1200):437–487, 1988.

[5] Roger T Hanlon and John B Messenger. *Cephalopod behaviour*. Cambridge University Press, 2018.

[6] Roger Hanlon, Mike Vecchione, and Louise Allcock. Octopus, squid, and cuttlefish: A visual, scientific guide to the oceans' most advanced invertebrates. 2018.

[7] Angel Guerra. Ecology of sepia officinalis. *Vie et Milieu/Life & Environment*, pages 97–107, 2006.

[8] Malcolm R Clarke and JD Stevens. Cephalopods, blue sharks and migration. *Journal of the marine biological association of the United Kingdom*, 54(4):949–957, 1974.

[9] Salome Morte, Manuel J Redon, and Antonio Sanz-Brau. Feeding habits of juvenile

mustelus mustelus (carcharhiniformes, triakidae) in the western mediterranean. *Cahiers de biologie marine*, 38(2):103–108, 1997.

[10] F Velasco, I Olaso, and F Sánchez. The role of cephalopods as forage for the demersal fish community in the southern bay of biscay. *Fisheries Research*, 52(1-2):65–77, 2001.

[11] MG Larrañeta. Sobre la alimentación, la madurez sexual y la talla de primera captura de merluccius merluccius (l.). 1970.

[12] Patrick Le Mao. Place de la seiche sepia officinalis (mollusque céphalopode) dans les chaines alimentaires du golfe normano-breton. *Cahiers de biologie marine*, 26(3):331–340, 1985.

[13] A Blanc and J Daguzan. Size selectivity in the diet of the young cuttlefish sepia officinalis (mollusca: Sepiidae). *Journal of the Marine Biological Association of the United Kingdom*, 80(6):1137–1138, 2000.

[14] Gerald Handerson Thayer. *Concealing-coloration in the animal kingdom: an exposition of the laws of disguise through color and pattern*. Macmillan Company, 1918.

[15] Hugh B Cott. Adaptive coloration in animals. 1940.

[16] Michelle D Staudinger, Kendra C Buresch, Lydia M Mäthger, Charlie Fry, Sarah McAnulty, Kimberly M Ulmer, and Roger T Hanlon. Defensive responses of cuttlefish to different teleost predators. *The Biological Bulletin*, 225(3):161–174, 2013.

[17] Martin Stevens and Sami Merilaita. Animal camouflage: current issues and new perspectives. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364 (1516):423–427, 2009.

[18] Sami Merilaita, H Martin Schaefer, and Marina Dimitrova. What is camouflage through distractive markings? *Behavioral Ecology*, 24(5):e1271–e1272, 2013.

[19] Kay M Cooper, Roger T Hanlon, and Bernd U Budelmann. Physiological color change in squid iridophores. *Cell and tissue research*, 259(1):15–24, 1990.

[20] Lydia M Mäthger. The response of squid and fish to changes in the angular distribution of light. *Journal of the Marine Biological Association of the United Kingdom*, 83(4): 849–856, 2003.

[21] LM Mathger and EJ Denton. Reflective properties of iridophores and fluorescent 'eyespots' in the loliginid squid alloteuthis subulata and loligo vulgaris. *Journal of Experimental Biology*, 204(12):2103–2118, 2001.

[22] BW Jones and MK Nishiguchi. Counterillumination in the hawaiian bobtail squid, euprymna scolopes berry (mollusca: Cephalopoda). *Marine Biology*, 144(6):1151–1155, 2004.

[23] Roger T Hanlon, Marie-Jose Naud, John W Forsythe, Karina Hall, Anya C Watson, and Joy McKechnie. Adaptable night camouflage by cuttlefish. *The American Naturalist*, 169(4):543–551, 2007.

[24] Roger T Hanlon, John W Forsythe, and David E Joneschild. Crypsis, conspicuousness, mimicry and polyphenism as antipredator defences of foraging octopuses on indo-pacific coral reefs, with a method of quantifying crypsis from video tapes. *Biological Journal of the Linnean Society*, 66(1):1–22, 1999.

[25] Alexandra Barbosa, Lydia M Mäthger, Kendra C Buresch, Jennifer Kelly, Charles Chubb, Chuan-Chin Chiao, and Roger T Hanlon. Cuttlefish camouflage: the effects of substrate contrast and size in evoking uniform, mottle or disruptive body patterns. *Vision research*, 48(10):1242–1253, 2008.

[26] Roger Hanlon, Chuan-Chin Chiao, Lydia Mathger, Kendra Buresch, A. Barbosa, J.J. Allen, L. Siemann, and Charles Chubb. Rapid adaptive camouflage in cephalopods.

Animal Camouflage: Mechanisms and Function, pages 145–163, 01 2011. doi: 10.1017/CBO9780511852053.009.

[27] Jolyon Troscianko, John Skelhorn, and Martin Stevens. Quantifying camouflage: how to predict detectability from appearance. *BMC evolutionary biology*, 17(1):1–13, 2017.

[28] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[29] Matthias Kummerer, Thomas SA Wallis, Leon A Gatys, and Matthias Bethge. Understanding low-and high-level contributions to fixation prediction. In *Proceedings of the IEEE international conference on computer vision*, pages 4789–4798, 2017.

[30] Matthias Kummerer, Thomas SA Wallis, and Matthias Bethge. Saliency benchmarking made easy: Separating models, maps and metrics. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 770–787, 2018.

[31] Matthias Kümmerer, Lucas Theis, and Matthias Bethge. Deep gaze i: Boosting saliency prediction with feature maps trained on imagenet. *arXiv preprint arXiv:1411.1045*, 2014.

[32] Leon A Gatys, Alexander S Ecker, and Matthias Bethge. A neural algorithm of artistic style. *arXiv preprint arXiv:1508.06576*, 2015.

[33] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database.

[34] Thomas Serre, Lior Wolf, Stanley Bileschi, Maximilian Riesenhuber, and Tomaso Poggio. Robust object recognition with cortex-like mechanisms. *IEEE transactions on pattern analysis and machine intelligence*, 29(3):411–426, 2007.

[35] Javier Portilla and Eero P Simoncelli. A parametric texture model based on joint statistics of complex wavelet coefficients. *International journal of computer vision*, 40 (1):49–70, 2000.

[36] B Julesz. A theory of preattentive texture discrimination based on first-order statistics of textons, biological cybernetics, vol. 41. 1962.

[37] Yan Xu, Zhipeng Jia, Yuqing Ai, Fang Zhang, Maode Lai, and Eric I-Chao Chang. Deep convolutional activation features for large scale brain tumor histopathology image classification and segmentation. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 947–951, 2015. doi: 10.1109/ICASSP.2015.7178109.

[38] Joseph A Lombardo, Matthew V Macellaio, Bing Liu, Stephanie E Palmer, and Leslie C Osborne. State dependence of stimulus-induced variability tuning in macaque mt. *PLOS Computational Biology*, 14(10):e1006527, 2018.

[39] Marlene R Cohen and John HR Maunsell. Attention improves performance primarily by reducing interneuronal correlations. *Nature neuroscience*, 12(12):1594–1600, 2009.

[40] Jude F Mitchell, Kristy A Sundberg, and John H Reynolds. Differential attention-dependent response modulation across cell classes in macaque visual area v4. *Neuron*, 55(1):131–141, 2007.

[41] Bastian Schledde, F Orlando Galashan, Magdalena Przybyla, Andreas K Kreiter, and Detlef Wegener. Task-specific, dimension-based attentional shaping of motion processing in monkey area mt. *Journal of neurophysiology*, 118(3):1542–1555, 2017.

[42] F Orlando Galashan, Hanna C Saßen, Andreas K Kreiter, and Detlef Wegener. Monkey area mt latencies to speed changes depend on attention and correlate with behavioral reaction times. *Neuron*, 78(4):740–750, 2013.

[43] Robert Niebergall, Paul S Khayat, Stefan Treue, and Julio C Martinez-Trujillo. Expansion of mt neurons excitatory receptive fields during covert attentive tracking. *Journal of Neuroscience*, 31(43):15499–15510, 2011.

[44] Umesh Rajashekar, Lawrence K Cormack, and Alan C Bovik. Visual search: Structure from noise. In *Proceedings of the 2002 symposium on Eye tracking research & applications*, pages 119–123, 2002.

[45] Wilson S Geisler, Jeffrey S Perry, and Jiri Najemnik. Visual search: The role of peripheral information measured using gaze-contingent displays. *Journal of Vision*, 6(9):1–1, 2006.

[46] Armin Schneider and Hubertus Feussner. Chapter 5 - diagnostic procedures. In Armin Schneider and Hubertus Feussner, editors, *Biomedical Engineering in Gastrointestinal Surgery*, pages 87–220. Academic Press, 2017. ISBN 978-0-12-803230-5. doi: https://doi.org/10.1016/B978-0-12-803230-5.00005-1. URL `https://www.sciencedirect.com/science/article/pii/B9780128032305000051`.

[47] F. Vasefi, N. MacKinnon, and D.L. Farkas. Chapter 16 - hyperspectral and multispectral imaging in dermatology. In Michael R. Hamblin, Pinar Avci, and Gaurav K. Gupta, editors, *Imaging in Dermatology*, pages 187–201. Academic Press, Boston, 2016. ISBN 978-0-12-802838-4. doi: https://doi.org/10.1016/B978-0-12-802838-4.00016-9. URL `https://www.sciencedirect.com/science/article/pii/B9780128028384000169`.

[48] Misha Vorobyev and Daniel Osorio. Receptor noise as a determinant of colour thresholds. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265 (1394):351–358, 1998.

[49] Hannah Markowitz. Sepia latimanus. URL `https://animaldiversity.org/accounts/Sepia_latimanus/`.

[50] Dianne Aglibot. Sepia apama. URL https://animaldiversity.org/accounts/Sepia_apama/.

[51] Ae Lin Compton and Laura Wiley. Sepia officinalis. URL https://animaldiversity.org/accounts/Sepia_officinalis/.

[52] A Blanc and J Daguzan. Young cuttlefish sepia officinalis (mollusca: Sepiidae) in the morbihan bay (south brittany, france): accessory prey of predators. *Journal of the Marine Biological Association of the United Kingdom*, 79(6):1133–1134, 1999.

[53] DG Stavenga, RP Smits, and BJ Hoenders. Simple exponential functions describing the absorbance bands of visual pigment spectra. *Vision research*, 33(8):1011–1017, 1993.

[54] Victor I Govardovskii, Nanna Fyhrquist, Tom Reuter, Dmitry G Kuzmin, and Kristian Donner. In search of the visual pigment template. *Visual neuroscience*, 17(4):509–528, 2000.

[55] Andrij Z Horodysky, Richard W Brill, Eric J Warrant, John A Musick, and Robert J Latour. Comparative visual function in four piscivorous fishes inhabiting chesapeake bay. *Journal of Experimental Biology*, 213(10):1751–1761, 2010.

[56] Julia Shand, JC Partridge, SN Archer, GW Potts, and JN Lythgoe. Spectral absorbance changes in the violet/blue sensitive cones of the juvenile pollack, pollachius pollachius. *Journal of Comparative Physiology A*, 163(5):699–703, 1988.

[57] Solomon Kullback and Richard A Leibler. On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86, 1951.

[58] Avi Ben-Simon, Ohad Ben-Shahar, Genadiy Vasserman, Mor Ben-Tov, and Ronen Segev. Visual acuity in the archerfish: Behavior, anatomy, and neurophysiology. *Journal of Vision*, 12(12):18–18, 11 2012. ISSN 1534-7362. doi: 10.1167/12.12.18. URL https://doi.org/10.1167/12.12.18.