Preliminary Study Towards a Fuzzy Model for Visual Attention

Anca Ralescu[1], **Isabelle Bloch**[2], and **Roberto Cesar**[3]

1. EECS Department, University of Cincinnati, ML 0030, Cincinnati, OH 45221, USA - anca.ralescu@uc.edu
2. Institut Mines Telecom, Telecom Paristech, CNRS LTCI, Paris, France - isabelle.bloch@telecom-paristech.fr
3. University of Sao Paulo, IME, Sao Paulo, Brazil - cesar@ime.usp.br

Abstract

Attention, in particular visual attention, has been a subject of studies in various disciplines, including cognitive science, experimental psychology, and computer vision. In cognitive science and experimental psychology the objective is to develop theories that can explain the attention phenomenon of cognition. In computer vision, the objective is to inform image understanding systems by hypotheses on the human visual attention. There is, however, very little influence of studies across these two disciplines. In a departure from this state of affairs, this study seeks to develop an algorithmic approach to visual attention as part of an image understanding system, by starting with a theory of visual attention put forward in experimental psychology. In the process, it will become useful to revise some of the concepts of this theory, in particular by adopting fuzzy set based representations and the necessary calculus for them.

1 Introduction

As subject of human cognition, attention has attracted a great interest from the fields of cognitive science and experimental psychology.

Visual attention is a wide field, largely addressed in the literature covering different aspects. Some works related to the present paper are briefly reviewed, without seeking at exhaustivity. One approach relies on Gestalt theory, and Gestalt and computer vision models are compared by (Desolneux, Moisan, and Morel 2003). Two sets of experiments for Gestalt detection methods are carried out and compared to computationally predicted results. Object size and noise are the two parameters taken into account in these experiments. The authors indicate that the qualitative thresholds predicted by the proposed computational approach of gestalt detection fit the human perception.

Another approach is purely computational and based on image information. An important review on visual attention modeling is presented by (Borji and Itti 2013). The important aspect of saliency-based attention is specifically addressed in this review. Nearly 65 models are reviewed and classified in a didactical taxonomy that helps clarifying the field. Visual saliency refers to a bottom-up phenomenon where some scene regions are detected as more prominent than others due to some visual features. There are different biological and computational approaches to model such phenomena. For instance, the center-surround hypothesis (a common issue for the analysis of receptive fields in the retina) is a classical model for bottom-up saliency (Gao, Mahadevan, and Vasconcelos 2008). In such settings, Gao and co-authors (Gao, Mahadevan, and Vasconcelos 2008) incorporate discriminant features and decision-theoretic model for saliency characterization. Saliency detection is important in many different imaging and vision applications (Yan et al. 2013; Yang et al. 2013). For instance, in medical imaging, saliency maps are useful to guide model-based image segmentation (Fouquier, Atif, and Bloch 2012), thus merging top-down and bottom-up approaches.

The mechanism of attention has been studied intensively in the field of psychology and cognitive science, (Kahneman 1973), (Treisman and Gelade 1980), (Treisman 1988), (Treisman 2014), (Humphreys 2014), (Bundesen, Habekost, and Kyllingsbæk 2005) (Bundesen, Vangkilde, and Petersen 2014). In this paper we focus on the theory of visual attention introduced in (Bundesen 1990), where visual recognition and attentional selection are considered as the task of *perceptual categorization*, basically deciding to which category an object or element of the visual field belongs.

Following the notation of (Bundesen 1990), throughout this paper, x is an input item, e.g. image or image region, or more generally an item to be categorized of recognized. The collection of all items x is denoted by S. A category is denoted by i and the collection of all categories is denoted by R. A category can stand for an ontological category (e.g., an object, or a scene), or for subsets in the range of a particular attribute (e.g., red for the attribute color). Regardless of the situation the conceptual treatment of categories and/or items is the same. $E(x, i)$ denotes the event/statement "x is in category i". When viewed as an event, one can talk about *its probability*; when viewed as a statement, one can talk about its truth or *its possibility*.

From this point on this paper is organized as follows: Section 2 contains a brief review of TVA concepts and mechanisms - filtering and pigeonholing. Section 3 presents the motivation for the introduction of fuzzy sets; the fuzzy mechanisms of filtering and pigeonholing. Conclusions and future research are in Section 4.

2 TVA concepts and mechanisms of attention

In this section, we review and comment the main concepts and modeling steps of the Theory of Visual Attention (TVA) by (Bundesen 1990).

2.1 Attentional Weight

One of the main concepts introduced in TVA is that of attentional weight defined as follows:

$$w(x) = \sum_{i \in R} \eta(x,i)\pi(i) \tag{1}$$

What are the possible interpretations of the quantities in Equation (1)? If $\eta(x,i)$ is interpreted as *the salience of x for category i*, then $w(x)$ could be interpreted as the *salience of x across the family of categories R*, averaged with respect to category pertinence. From the point of view of computer vision, $\eta(x,i)$ is simply the output of an operator designed to provide information for category i.

Note that pertinence of a category is (or must be) considered with respect to a task, which could be a categorization at a higher semantic/ontological level. Adopting this point of view, the product $\eta(x,i)\pi(i)$ can then be interpreted as *the pertinence of item x to the task with respect to which category i had pertinence $\pi(i)$*. More precisely, one can define

$$\pi(x,T_i) = \eta(x,i)\pi(i)$$

as the pertinence of x to T_i where T_i is the task to which category i has pertinence value $\pi(i)$.

For example, suppose that i is the color category "red" of the attribute color. Furthermore, suppose that the color category "red" has pertinence $\pi(red)$ to the task of identifying visually an object such as, for instance, the "flag of some country". Let now x be a region in an image, and $\eta(x,red)$ the output of evaluating it with respect to the color "red". Then $\eta(x,T_{red}) = \eta(x,red)\pi(red)$ is the pertinence of x to the task T_{red}.

Taking max/min with respect to x obtains:

$$x_{max,red} = arg \max_{x \in S} \eta(x,T_{red}),$$

the region in the input which is most pertinent to T_{red}, and

$$x_{min,red} = arg \min_{x \in S} \eta(x,T_{red}),$$

the region in the input which is least pertinent to T_{red}.

Similarly, taking max / min over categories, yields

$$i_{max} = arg \max_{i \in R} \pi(i); \; i_{min} = arg \min_{i \in R, \pi(i)>0} \pi(i)$$

the most/least pertinent categories respectively. The condition $\pi(i) > 0$ ensures that categories which are not pertinent at all, i.e. with $\pi(i) = 0$, are not taken into account, so the trivial case $\pi(i_{min}) = 0$ is never obtained. Then, for fixed x, $\eta(x,i_{max})$, $\eta(x,i_{min})$ are the strengths of evidence for x to be in the highest/lowest pertinence category, and

$$\pi(x,T_{max}) = \eta(x,i_{max})\pi(i_{max})$$
$$\pi(x,i_{min}) = \eta(x,T_{min})\pi(i_{min})$$

are the importance of x to the task corresponding to the category of highest/lowest pertinence value. Versions of the following "flag example" will be used in this paper to illustrate various points.

Example 1 *Let T stand for the task to determine if an object identified in an image corresponds to a "flag of some country". The decision is to be based on color information only. Assume several color categories and their respective pertinences as shown in Table 1.*

Table 1: Color categories and their respective pertinence values to the task "Identify flag of a country".

Color category: i	Category pertinence: $\pi(i)$
red	0.8
yellow	0.3
black	0.1
green	0.2
$(\max \pi(i), i_{max})$	$(0.8, red)$
$(\min \pi(i), i_{min})$	$(0.1, black)$

In this example

$$\eta(x,T_{red}) = 0.8\eta(x,red); \; \eta(x,T_{black}) = 0.1\eta(x,black).$$

In Equation (1) only those categories i with $\pi(i) > 0$ contribute to $w(x)$. This means that categories which are not pertinent (i.e., $\pi(i) = 0$) are never considered for x, even when $\eta(x,i)$ is very large.

To summarize, with the interpretation of $\eta(x,i)\pi(i)$ as described above, the *attentional weight $w(x)$* defined by Equation (1) is the cumulative pertinence of x to a task T, obtained from *strength of the sensory evidence given by x to all categories, in proportion to their pertinence to the task T*.

2.2 Hazard Function

In (Bundesen 1990) the notion of a hazard function $\nu(x,i)$ is introduced as $\nu(x,i) = Prob(E(x,i))$, that is, the probability that item x is in category i (e.g., image region x is *red*). It is assumed (see 2nd assumption in (Bundesen 1990)) that ν is computed as:

$$\nu(x,i) = \eta(x,i)\beta(i)w(x) \tag{2}$$

where $\eta(x,i)$ and $w(x)$ are as described above[1], and $\beta(i)$ is introduced to indicate a *bias for category i*. Since ν is interpreted as a probability, $\nu(x,i) \in [0,1]$, which is ensured when $\eta(x,i), \beta(i), w(x) \in [0,1]$, without additional constraints on these values. Moreover, when R is an exhaustive set of exclusive (non-overlapping) categories, then ν should be normalized so that $\sum_{i \in R} \nu(x,i) = 1$, in order to really satisfy its interpretation from (Bundesen 1990) as a probability. More recently, in (Bundesen, Vangkilde, and Petersen 2014) $\beta(i)$ is decomposed as

$$\beta(i) = Ap(i)u(i) \tag{3}$$

where $A \in [0,1]$ is the level of alertness, and $p(i)$ and $u(i)$ are respectively, the prior probability and utility of category i. One can imagine that A also varies with the category, in

---

[1]Note that the expression of (Bundesen 1990) involves a normalized version of w, i.e. $w(x)/\sum_{x \in S} w(x)$. Here we implicitly assume that w is normalized, in order to simplify equations.

which case A in Equation (3) is replaced by an A_i. This is justified by the fact that one may be more alert to a category than to others. In an image processing system, A, or A_i could be tied to the performance of the image processing operators used. The components $p(i)$, $u(i)$ of $\beta(i)$, and hence $\beta(i)$, must also be tied up to a (higher level) task T. While $p(i)$ may be obtained from past data and experiments on the task T, $u(i)$ seems to be purely *subjective*, and to a large extent, its role seems to overlap with that is $\pi(i)$. Plugging $w(x)$ and $\beta(i)$ in (2) results in

$$
\begin{aligned}
\nu(x,i) \;\; &= A\eta(x,i)p(i)u(i)\sum_{j\in R}\eta(x,j)\pi(j)\\
&= Ap(i)u(i)[\eta(x,i)^2\pi(i)+\\
&\quad +\eta(x,i)\sum_{j\neq i}\eta(x,j)\pi(j)]
\end{aligned}
\tag{4}
$$

which suggests that the most important role in computing $\nu(x,i)$ is played by the sensory evidence. In particular, ν's largest value is obtained when $A = p(i) = u(i) = 1$, (i.e. under maximum alertness, maximum prior probability, and maximum utility), and in that case $\nu(x,i)$ is a function only of the sensory evidence. Stated differently, this means that A, $p(i)$ and $u(i)$ can only decrease the value of $\nu(x,i)$. However, they may provide a mechanism to account for different types of subjective information, and of ranking the values of $\nu(x,i)$ when they enter its definition as shown in Equations (2) - (4). The justification in (Bundesen, Vangkilde, and Petersen 2014) of Equation (3) is based on the fact that when either one of A, $p(i)$, or $u(i)$ is null, then $\beta(i) = 0$. However, the same result holds when these quantities enter the definition of β not through a product, but through other operations, such as the $\min$, or more generally, t-norms.

The fact that the value of $\nu(x,i)$ decreases when $Ap(i)u(i) \neq 1$ (i.e. at least one of these three values is less than 1, $u(i)$ for instance) can be interpreted as follows: x will be less probably categorized in i if, for instance, the utility for i is low, which means that we do not really care for this category. This also goes with the interpretation as a rate of encoding information in the memory, according to (Bundesen 1990), even without considering time information.

The two mechanisms for visual attention proposed in (Bundesen 1990), filtering and pigeonholing, are described next.

2.3 Filtering

Filtering (Bundesen 1990) refers to the mechanism of selecting an item $x \in S$ (given a higher level task), for a target category i. This mechanism seeks to

(F1) increase $\nu(x,i)$ for some category i, while

(F2) not changing the conditional probability of $E(x,i)$ given that x is categorized.

Filtering can be achieved by increasing $w(x)$ as follows:

For category $j \in R$ assume $\pi'(j) = a\pi(j)$, where $a > 1$. Then $w(x)$ of equation (1) becomes $w'(x) = \sum_{i\in R, i\neq j}\eta(x,i)\pi_i + \eta(x,j)\pi'_j = \sum_{i\in R, i\neq j}\eta(x,i)\pi_i + \eta(x,j)a\pi_j > w(x)$. Therefore, $\nu(x,i)$ becomes $\nu'(x,i) = \eta(x,i)\beta(i)w'(x) > \nu(x,i)$, which satisfies condition (F1)

above. Computing now $P(x \text{ is } i \mid x \text{ is categorized})$ yields:

$$
\begin{aligned}
P(x \text{ is } i \mid x \text{ is categorized}) &= \frac{\nu(x,i)}{\sum_{k\in R}\nu(x,k)}\\
= \frac{\eta(x,i)\beta(i)w(x)}{w(x)\sum_{k\in R}\nu(x,k)} &= \frac{\eta(x,i)\beta(i)}{\sum_{k\in R}\nu(x,k)}
\end{aligned}
\tag{5}
$$

which does not depend on w, hence satifies condition (F2). In Equation (5) the numerator is $\nu(x,i)$ since

$$
\{x \text{ is } i\} \subset \{x \text{ is categorized}\}
$$

and therefore

$P(x \text{ is } i, x \text{ is categorized}) = P(x \text{ is } i)$, while the denominator uses an assumption on non-overlapping categories to write $P(x \text{ is categorized})$ as $\sum_{k\in R}\nu(x,k)$. Dropping the constraint of non-overlapping categories is discussed later in this study.

2.4 Pigeonholing

For fixed item $x \in S$, pigeonholing (Bundesen 1990) refers to the mechanism of selecting a category $i \in R$ (given a higher level task), across a set of items S. It seeks to:

(P1) increase $\sum_{x\in S}\nu(x,i)$ for category i pertinent to the task, such that

(P2) for all $j \in R$, $j \neq i$, $\sum_{x\in S}\nu(x,j)$ does not change

Pigeonholing can be done by increasing $\beta(i)$ for some $i \in R$ as follows: For category $i \in R$, let $\beta'_i = a\beta_i$, with $a > 1$. Then

$$
\begin{aligned}
\nu'(x,i) \;\; &= \eta(x,i)\beta'_i w_x = \eta(x,i)a\beta_i w_x\\
&> \eta(x,i)\beta_i w_x = \nu(x,i).
\end{aligned}
$$

Summing up over $x \in S$ obtains

$$
P'(i \text{ is selected}) = \sum_{x\in R}\eta(x,i)\beta'_i w_x > P(i \text{ is selected}),
\tag{6}
$$

which achieves (P1). At the same time, it is clear that for any other category $j \neq i$, $P(j \text{ is selected})$ does not change, and hence (P2) is satisfied too.

Equation (6) uses the assumption that items x are non-overlapping, for example that they form a partition of the image. However, this partition need not be crisp, i.e. may allow overlapping x's, as for example these are stated in qualitative terms. In such cases, Equation (6) does not hold. Dropping the constraint of non-overlapping items, discussed later, leads to a different interpretation of $\nu(x,i)$.

3 Fuzzy Mechanisms for Visual Attention

We consider in this section the situations when the values of the attentional weight and/or category pertinence are not exact. In such situations these values may be represented as fuzzy sets, and therefore, the computation of the categorization of an item must resort to calculus with fuzzy sets. First, let us see why indeed such situations may arise.

Recall that in its original definition, for a given input x and category i, the *strength of sensory evidence for $E(x,i)$*, $\eta(x,i) \in [0,1]$. Assuming that $\eta(x,i)$ is the output of an operator/test for category i on item x, this output may be inexact because of the inexact nature of the category i. For

example, if the category $i = red$ of the attribute color, then for a given input pixel value x this category holds "more or less" and it may not be useful to commit to an exact 0/1 value.

Likewise, in its original definition, the pertinence of a category, $\pi(i)$ conveys its importance. Obviously, given a collection of visual categories, and task, they may be distinguished along their pertinence values. Moreover, several categories may have the same, maximum importance for the given task. As an example, consider the pertinence of color categories for the detection of an object which is known to have one of two possible color categories, white or yellow, from the collection of all possible color categories. In this case, it is useful to be able to encode

$$\pi(white) = \pi(yellow) = 1,$$

which would be possible when π is considered as a *possibility distribution* on the color categories, regardless of the number of color categories allowed. By contrast, using a probability based approach, the cardinality of R, the collection of categories, restricts the values assigned to these equally possible categories, to at most 0.5. That is, $\pi(white) = \pi(yellow) \leq 0.5$ with equality when $R = \{yellow, white\}$.

3.1 A new definition for $w(x)$

The departure point for the new definition for $w(x)$ is the interpretation of a special case of Equation (1). Let $R_a = \{i \in R \mid \pi(i) = a\}$ and consider the special case $R = R_0 \cup R_1$, that is, all categories in R are either "fully" pertinent, $\pi(i) = 1$ ($i \in R_1$), or not pertinent $\pi(i) = 0$ ($i \in R_0$). Then (1) becomes

$$w(x) = \sum_{i \in R_1} \eta(x, i)$$

Next let $\eta_{max} = \max_{i \in R_1}$, and recall that $\eta(x, i) \leq 1$. Then

$$w(x) \leq \sum_{i:\pi(i)=1} \eta_{max} = \eta_{max} \sum_{i \in R_1} 1 = \eta_{max}|R_1| \leq |R_1|,$$

where $|R_1|$ denotes the cardinality of the set R_1. That is, $w(x)$ is bounded by the number of categories i with pertinence $\pi(i) = 1$. If $\eta(x, i) = 1$ for all $i \in R_1$ then $w(x)$ is exactly the number of such categories.

This meaning of $w(x)$ is very natural and appealing. Indeed, one would expect the item x to count to the extent that it supports more categories. To generalize this notion, define for fixed $x \in S$ and fixed task T

$$\mu_{(x,T)}(i) = \eta(x, i)\pi_T(i)$$

the *degree* to which category i, pertinent to task T, is supported by the (data) item x as shown by the strength of sensory evidence, $\eta(x, i)$. Therefore, $\mu_{(x,T)} : R \to [0, 1]$ is the membership of a fuzzy set on the set of categories. [2] Then the weight of item x is now defined as the cardinality of this fuzzy set. That is

$$\widetilde{w}(x) = Card\{(i, \mu_x(i)) \mid i \in R\} \qquad (7)$$

---

[2] In the following, assuming only one task, T, for ease of notation, the subscript T will be dropped, to write $\mu_x(i)$.

Several formulas for the cardinality of a fuzzy set have been put forward. Here, for illustration purposes, the definition from (Ralescu 1986) is used to obtain

$$Card\left(\{\mu_x(i) \mid i \in R\}\right)(k) = \mu_{x,(k)} \wedge (1 - \mu_{x,(k+1)}) \quad (8)$$

where $\mu_{x,(k)}$ denotes, the kth largest value of $\mu_x(\cdot)$, and $\mu_{x,(|R|+1)} = 0$. Thus, the cardinality defined in Equation (7) is a fuzzy set on $\{0, ..., |R|\}$. For an exact value of $w(x)$ the 0.5-level set of $\widetilde{w}(x)$ (which is an interval), or its classic cardinality can be used (Ralescu 1995).

3.2 A new definition for $\beta(i)$

Following the discussion from Section 2.4, define

$$\widetilde{\beta}(i) = \min\{A, p(i), u(i)\} \qquad (9)$$

As in the case of β defined in (3), $\widetilde{\beta}(i) = 0$ whenever $A = 0$, or $p(i) = 0$, or $u(i) = 0$, and the discussion of (Bundesen 1990) holds: that is, category i biases the selection to the extent that the system is alert, and category i is possible and useful. Alternatively, (9) means that the bias for the selection of i cannot be greater than the system alertness, the possibility of i or its utility. Furthermore, replacing the product by $\min$ also eliminates the possibility of values for $\widetilde{\beta}$ smaller than each one of A, $p(i)$, and $u(i)$, which is the well-known drowning effect of multiplication of positive values smaller than 1. More importantly, it should be mentioned that the $\min$ can handle *ordinal* or *qualitative values*, without needing specifying precise numbers. Specifying such precise values might be difficult when subjective assessments are made. By contrast, in the case of such assessments, ordinal or qualitative values are usually easily produced.

As already mentioned, in the fuzzy set framework, the product and $\min$ are but two particular cases of a t-norm (conjunction operator). A, $p(i)$, and $u(i)$ are interpreted respectively, as degrees of alertness, possibility (rather than probability) of i to be selected, and utility for the category i, and the bias for i is defined as the conjunction of these. This interpretation makes (9) meaningful beyond a mere computational artifice. Another choice for defining $\widetilde{\beta}$ is to select a more general, aggregation operator, $H : [0, 1] \times [0, 1] \times [0, 1] \to [0, 1]$, which would allow the contribution of more than one of A, $p(i)$, $u(i)$ towards $\widetilde{\beta}$.

3.3 A new definition for $\nu(x, i)$

With the new definitions, $\widetilde{w}(x)$, and $\widetilde{\beta}$ of $w(x)$ and β respectively, the meaning of $\nu(x, i)$ also changes from a probability to a possibility, more precisely, $Possibility(x \text{ is } i)$:

$$Possibility(x \text{ is } i) = H(\eta(x, i), \widetilde{\beta}(i), \widetilde{w}(x)) \qquad (10)$$

where H is again an aggregation operator, and hence the definition of $\nu(x, i)$ from (Bundesen 1990) is a particular case, when H is the product.

For defining H, one may rely on the huge literature on information fusion, for which the fuzzy sets theory provides a number of useful operators (see e.g. (Dubois and Prade 1985; Yager 1991; Bloch 1996) for reviews on fuzzy fusion operators). The large choice offered by these operators

allows modeling different combination behaviors (conjunctive, disjunctive, compromise, etc.), with different degrees (e.g. the min is a less severe conjunction as the product). Operators can also behave differently depending on whether the values to be combined are small, large, of the same order of magnitude, or having different priorities. The operators H could also be set differently for the three values. For instance η and $\widetilde{w}$, which depend on x and i could be combined using an operator H_1, and the result combined with $\widetilde{\beta}$, which depends on i only, using another operators H_2.

4 Conclusions and Future Work

This paper discussed an attentional model developed in the field of psychology and cognitive science set in a probabilistic framework. The basic concepts of this model were discussed and an alternative, fuzzy set based approach was suggested. In the fuzzy set framework, modeling would be easier, more natural (for instance replacing numbers by ordinal or qualitative values), and it would allow for more flexible ways of combining the different terms. This discussion paves the way for a new attentional model, the complete development of it being left for future work.

5 Acknowledgments

References

Bloch, I. 1996. Information Combination Operators for Data Fusion: A Comparative Review with Classification. *IEEE Transactions on Systems, Man, and Cybernetics* 26(1):52–67.

Borji, A., and Itti, L. 2013. State-of-the-art in visual attention modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35(1):185–207.

Bundesen, C.; Habekost, T.; and Kyllingsbæk, S. 2005. A neural theory of visual attention: bridging cognition and neurophysiology. *Psychological review* 112(2):291.

Bundesen, C.; Vangkilde, S.; and Petersen, A. 2014. Recent developments in a computational theory of visual attention (tva). *Vision research*.

Bundesen, C. 1990. A theory of visual attention. *Psychological review* 97(4):523.

Desolneux, A.; Moisan, L.; and Morel, J.-M. 2003. Computational gestalts and perception thresholds. *Journal of Physiology-Paris* 97(2):311–324.

Dubois, D., and Prade, H. 1985. A Review of Fuzzy Set Aggregation Connectives. *Information Sciences* 36:85–121.

Fouquier, G.; Atif, J.; and Bloch, I. 2012. Sequential model-based segmentation and recognition of image structures driven by visual features and spatial relations. *Computer Vision and Image Understanding* 116(1):146–165.

Gao, D.; Mahadevan, V.; and Vasconcelos, N. 2008. The discriminant center-surround hypothesis for bottom-up saliency. In *Advances in Neural Information Processing Systems*, 497–504.

Humphreys, G. W. 2014. Feature confirmation in object perception: Feature integration theory 26 years on from the Treisman Bartlett lecture. *The Quarterly Journal of Experimental Psychology* (just-accepted):1–49.

Kahneman, D. 1973. *Attention and Effort*. Prentice-Hall.

Ralescu, A. L. 1986. A note on rule representation in expert systems. *Information Sciences* 38(2):193–203.

Ralescu, D. 1995. Cardinality, quantifiers, and the aggregation of fuzzy criteria. *Fuzzy sets and systems* 69(3):355–365.

Treisman, A. M., and Gelade, G. 1980. A feature-integration theory of attention. *Cognitive psychology* 12(1):97–136.

Treisman, A. 1988. Features and objects: The fourteenth bartlett memorial lecture. *The Quarterly Journal of Experimental Psychology* 40(2):201–237.

Treisman, A. 2014. The psychological reality of levels of processing. *Levels of processing in human memory* 301–330.

Yager, R. R. 1991. Connectives and Quantifiers in Fuzzy Sets. *Fuzzy Sets and Systems* 40:39–75.

Yan, Q.; Xu, L.; Shi, J.; and Jia, J. 2013. Hierarchical saliency detection. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1155–1162.

Yang, C.; Zhang, L.; Lu, H.; Ruan, X.; and Yang, M.-H. 2013. Saliency detection via graph-based manifold ranking. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 3166–3173.