

# Clothing and People - A Social Signal Processing Perspective

Maedeh Aghaei<sup>1</sup>, Federico Parezzan<sup>3</sup>, Mariella Dimiccoli<sup>1</sup>, Petia Radeva<sup>1,2</sup>, Marco Cristani<sup>3</sup>

<sup>1</sup> Faculty of Mathematics and Computer Science, University of Barcelona, Barcelona, Spain

<sup>2</sup> Computer Vision Center, Barcelona, Spain

<sup>3</sup> Department of Computer Science, University of Verona, Verona, Italy

**Abstract**—In our society and century, clothing is not anymore used only as a means for body protection. Our paper builds upon the evidence, studied within the social sciences, that clothing brings a clear communicative message in terms of social signals, influencing the impression and behaviour of others towards a person. In fact, clothing correlates with personality traits, both in terms of self-assessment and assessments that unacquainted people give to an individual. The consequences of these facts are important: the influence of clothing on the decision making of individuals has been investigated in the literature, showing that it represents a discriminative factor to differentiate among diverse groups of people. Unfortunately, this has been observed after cumbersome and expensive manual annotations, on very restricted populations, limiting the scope of the resulting claims. With this position paper, we want to sketch the main steps of the very first systematic analysis, driven by social signal processing techniques, of the relationship between clothing and social signals, both sent and perceived. Thanks to human parsing technologies, which exhibit high robustness owing to deep learning architectures, we are now capable to isolate visual patterns characterising a large types of garments. These algorithms will be used to capture statistical relations on a large corpus of evidence to confirm the sociological findings and to go beyond the state of the art.

## I. INTRODUCTION

Despite rich body of work on the role of behavioral cues in non-verbal social signal processing [43], deeper understanding of social signals requires further cues discovery. In this respect, some visual behavioral cues such as gesture, posture, gaze, physical appearance, and proxemics have received high attention, while other possible features such as clothing have been traditionally little studied [42].

This is an important lack in the social signal processing literature, since clothing affects behavioral responses in the form of impression formation or person self-perception. Several past studies in the social sciences aimed to assess this influence, showing that formality of the clothing influences impression of others towards a person [10] as well as the self-perception of people towards themselves [13], [1]. More recently, the influence of clothing on the decision making of individuals has been investigated [38]. A few studies also have shown that clothing may be an indicator of ethnicity, culture, socioeconomic status [39], [19], and even surroundings of the people [26].

This work was partially founded by projects TIN2012-38187- C03-01 and SGR 1219 and joint project 2016 grant program, University of Verona. M.A. is supported by an APIF grant from the University of Barcelona. P.R. is partly supported by an ICREA Academia grant.

978-1-5090-4023-0/17/\$31.00 ©2017 IEEE

Other studies shows that clothing correlates with the personality traits of people in a way that people with formal clothing perceive actions and objects, the inter-relationship and the intra-relationship between them in a more meaningful manner [36]. Clothing can make a person feel comfortable or not in a social situation [16] and can be considered as a determinant of how long it takes for strangers to trust one and how much they may trust them [40]. Aforementioned studies are evidences of the importance of clothing in social signaling. Arguably, clothing can be considered as the most evident blueprint of individuals, which is completely dependent on their conscious choices, is not as transient as a gesture, and is more evident than any micro-signals such as a sarcastic smile among the facial expressions.

Various experiments have been previously performed to measure the clothing effect on human behavior. According to the critical review of Johnson *et al.* [18], the effect of clothing on human behavior, usually is measured in combination with other variables. However, despite the rich body of work, to the best of our knowledge, all of the previous experimental studies were performed and analyzed manually.

In this position paper, we will sketch the future steps of the first systematic study on which social signals are conveyed by clothing, proposing a framework within the scope of computer vision to measure the clothing effect on the impression that we have on ourselves and that we trigger in the others. More precisely, in a first phase we will investigate the basic visual cues that could be associated to social signals; for example, checking how much tight shirts are associated to the social signal of attractiveness. In a second phase, we will perform a higher level analysis, investigating the types of behaviors that may result from a social interaction, in dependence on the type of garments worn by the interacting people; for example, analyzing interactions between formally versus casually suited individuals. All of this would be possible since computer vision technologies are now mature for a fine-grained analysis of the clothing, providing precise dense segmentations of outfits as results of human parsing algorithms, and automatically recognizing diverse clothing items [44], [46], [8] and styles [20].

The paper is organized as follows: in Sec. II the literature on clothing in terms of social sciences is reviewed; in Sec. III, clothing analysis in terms of human parsing approaches is reported. The core of the paper is Sec. IV, where we discuss our ideas related to the study of clothing under the social signal processing umbrella. The paper ends

with some final remarks in Sec. V.

## II. CLOTHING AND SOCIAL SEMIOTICS

Semiotics, as originally defined by Ferdinand de Saussure, is “the science of the life of signs in society” [6]. Semiotics investigates *signs* and analyzes them to provide significance to a specific problem. There are three main elements in semiotics: the sign, what it refers to, and the people who use it. The people as social species and biological entities, instinctively evolved to survive better through facilitating living in a disciplined society by defining new signs and giving them an appropriate interpretation. Social semiotics is a subcategory of semiotics that studies how people design and interpret meanings and how these meanings are shaped by a specific social situation [15]. In social semiotics, the term *resource* is preferred over the term *sign* and represents a used signifier by the people to produce and to interpret communicative artefacts. In this respect, social semiotics is particularly useful in disclosing unnoticed significance and functionality of social resources and each individual is a semiotician, since everybody constantly interprets the meaning of signs around them.

Humans signify specific social context through *resources* of all type, whether visual, verbal or gestural. *Clothing* is a non-verbal resource that transfers meanings about individuals in the society. Cloths hold a symbolic and communicative role having the capacity to express style, identity, profession, social status, and gender or group affiliation of an individual. Although the symbolism that clothing carries on is not always clear, it evidently can be considered as the most desired personal image that one is willing to project to the society [29]. The study of how people use and interpret specific social context through dress is known as *clothing semiotics* or *fashion semiotics*, although, as stated by Sproles and Burns, clothing is distinct from fashion [37]. In this definition, clothing is “any covering for the human body”, while fashion is “the style of dress that is temporarily adopted by a discernible proportion of members of a social group, because that chosen style is perceived to be socially appropriate for the time and situation”. Originally, clothing semiotics was studied from the fine arts perspective. Later, the perspective has been expanded and the study covered the human needs in this respect [31]. Subsequently, in the 1960s, the social and psychological implications of clothing began receiving more attention from scholars. Today, clothing remains a common topic of study in social psychology [28] as it conveys social meaning about an individual and groups of people. It is in this way that the semiotics of clothing can be linked to the social semiotics.

In spite of the fact that clothes have such large potential to convey a message, it must be noted that clothing semiotics understanding is complex. The social context affects the interpretation of clothing, thus, having a precise knowledge of the unconscious symbolism attached to forms, colors, textures, postures, and other expressive elements that affects the interpretation of clothing in a given culture is a desired quality in automatic analysis of this information.

## III. CLOTHING AND COMPUTER VISION

*Clothing style* is commonly intended in computer vision as the set of visual attributes and category labels that describe an outfit [4], [46]. Examples of visual attributes are *colors* (red, green, etc.), *clothing patterns* (solid, striped, etc.), and more technical *qualitative expressions* (skin exposure, placket presence, etc.) [4]. These visual attributes have been used to measure the similarity between outfits paving the road for the identification and analysis of visual trends in fashion [44]. The category labels are textual expressions that individuate a particular type of clothing item (shirt, sweater, etc.) [46]. In most of the cases, all these textual labels are accompanied by a pixel-wise segmentation of the outfit, in which each segment is associated to a category label, and to one [23] or more visual attributes [7] (see Fig. 1).

This segmentation is the output of an operation commonly referred as *human* [9], [24], [22], *clothing* [46] or *fashion parsing* [23]. Clothing style can be also modeled without referring to a particular outfit, but instead to a larger set of category labels and visual attributes [20].

Human parsing is usually performed by statistical classifiers, which operate after a training phase. The training data may consist of fully labeled data, which means images of individual outfits in which each of the pixel has a label indicating the category and/or the attribute [46]. This is the most reliable source of information to train a classifier, but it is extremely cumbersome to get: in facts, manual annotation is necessary, which requires 15-60 minutes to be carried out for each single image [21]. Alternatively, weak labeling can be provided, which means to have training data in which an entire image is associated to a set of textual labels (in other words, the textual labels are not localized over the pixels of the outfit image). This obviously reduces the human labor to get training samples, but at the same time is less expressive, leading to classifiers which are not completely automatic: for example, [23] requires that the testing image too comes with textual labels that indicate what to look for in the image.

Most of the techniques for human parsing builds upon a preliminary operation, which is that of fitting a skeleton on the human body depicted in the input image. This operation is called pose estimation [47], and helps to introduce a structural prior for the parsing process, which individuate salient joints (ankles, knees, hips, shoulders, elbows, wrist, neck). These points are connected by sticks forming a skeleton, which in turn drive the parsing to align with it, providing anatomically plausible segmentations [46]. Unfortunately, pose estimation techniques are prone to errors in the case of missing data, due to occlusions or auto-occlusions; for this reason, images of single persons where the entire body is portrayed, are preferred. Images depicting parts of the body (as those ones captured via wearable sensors, where usually the whole body does not fit) represent a serious issue. In addition, pose estimation is weak in the case of large and long clothing, covering the structure of the body for what concerns some of the joints (a person wearing long dress has its knees completely covered). This issue has been recently



Fig. 1. Parsing example. The input image (left) and the final output of parsing (right).

faced by facing human parsing and pose estimation as two intertwined aspects of the same problem, introducing the concept of *semantic part* (such as leg, arm, head) [9]. A semantic part can be iteratively modeled with tools usually employed for human parsing (as the *Parselets* [8]) and as an ensemble of joints, taking from the pose estimation literature.

#### IV. CLOTHING STYLE AND SOCIAL SIGNALS: THE SOCIAL SIGNAL PROCESSING POINT OF VIEW

Social signal processing (SSP) “aims at providing computers with the ability to sense and understand human social signals” [42]. To this aim, Vinciarelli et al. [42] suggest a four-step pipeline in which, after having recorded the scene and detected humans (step 1 and 2), in the step 3, feature extraction has to be performed, where features are behavioral cues whose interpretation brings to individuate social signals. In the step 4, social signals have to be grounded with the scene context, in order to understand social interactions. We are interested specifically in the last two steps of the pipeline, since we assume that the scene has been already recorded and the individuals have been properly detected.

In the rest of this section, we will individuate the research questions (indicated with the letter **Q**) that can be inserted in these two steps, providing our intuitions about possible answers (letter **A**), driven by the literature of the human sciences and/or our speculations, together with the type of experiments we would like to carry out, to provide the community with deeper insight and novel tools for clothing social signaling.

##### A. Clothing behavioral cues as an individual social signal

**A-Q1 - How much clothing-related cues are independent from other standard behavioral cues, in the determination of particular social signals?** The question essentially asks how the mapping from visual features related to clothing (for example, the type and appearance of a particular clothing item, e.g., a shirt) has to be carried out in dependence from other cues such as the ones reported in [42] (Table 1, pag.1745). In other words, this question is very preliminary and asks for a feasible and reliable protocol with which clothing-related cues can be analyzed without caring of the effects due to other features in determining social signals.

**A-A1** - In social situations a clothing outfit comes with the body that wears it, so that other cues, in particular related to physical appearances, gesture and posture, face, emotional expressions and eyes behavior [42], [11] are obviously co-present and some cues may have different effects depending on the visible human body. For example, facial expression comes more into vision if only the upper body is visible. This could be the reason why online shops often present garments without the human body (Fig. 2). Thus, an analysis on these data seems reasonable and may help for answering A-Q1.

A simple yet important experiment would be that of checking whether the presence of different types of body appearances will change the nature of the social signal transmitted. In particular, our first step is to enrich the annotation of a clothing dataset, for example, the *Exact Street2Shop* dataset [12]. For a given garment, the dataset contains some “shop pictures”, where the garment is usually located on a neutral background, without being worn by a human body. Together with this, the dataset offers a “street picture”, where the same garment is worn by a subject among an undefined set of people. The idea is to first annotate standard semantic information about the people in the street photos (gender, expressions etc.). Next, different assessors will evaluate the street and the shop photos, defining the person wearing that particular garment in terms of social signals and personality traits. In the case of the street photos, the person in the picture is present and annotated, while in the case of the shop photos, persons are absent. The goal is to discover whether the presence of the person changes significantly the judgment of the assessors, and if this correlates with the semantic information associated to the person.

A finer setup, given a particular person, could be that of isolating the most the cues related to clothing by masking behavioral cues coming from the face (blurring the face) or hiding the height (removing the background scene). The interrelation between behavioral cues and other features in terms of social signals has never been investigated in the literature.

**A-Q2 - How to evaluate the nature of a social signal generated by clothing behavioral cues?** For understanding this question, one may consider the *Brunswick’s Lens* model. A simplified version of this model is adopted by Cristani et



Fig. 2. The picture shows an example of clothing outfits typical of online shops.

al. [5] (see Fig. 3). In few words, the model says that a social signal is not necessarily univocally intended. More in the detail, the model assumes that a social signal is sent by a sender,  $S$ , as a consequence of its internal state,  $\mu_s$ , which is assumed to be measurable. For example,  $S$  feels himself extrovert (his internal state), and this awareness is measured by a self-assessment (for example, using the Big Five questionnaire [3]).  $S$  wears some clothing items and as we are assuming that clothing items are related in some way with the internal state of  $S$ , they can be assumed as an *externalization* of the internal state. The receiver  $R$  sees the clothing items worn by  $S$ , and infers about the internal state of  $S$ , which in this case is  $\mu_r$ , to highlight that possibly is not equal to  $\mu_s$ . This process, called *attribution*, which brings to a perceived state what can be measured itself. The Brunswick Lens model states that a social signal has high ecological validity  $\rho_{EV}$ , if there is a high correlation between the internal state of  $S$  and the features. Viceversa, a social signal has high representational validity  $\rho_{RV}$ , if the correlation between the features and the state inferred by  $R$  is high. Finally, if the internal state of  $S$  correlates with the inferred state of  $R$ , it means that the communication through the social signal mediated by the features has high functional validity  $\rho_{FV}$ .

**A-A2** - The Brunswick's Lens essentially states that the nature of a social signal should be measured considering the sender of the signal and the receiver. This opens up to diverse experiments, suggesting a protocol for each one of them. For example, in order to understand how a particular social signal built upon clothing behavioral cues is interpreted by a generic receiver, it is necessary to measure the perceived state of multiple assessors. If the correlation between the behavioral cues and the perceived state of the assessors is high, we may individuate the *implied* meaning of a particular outfit. In more practical terms, to assess whether an athletic outfit is a behavioral cue that communicates the social signal of extroversion, this can be asked to a set of assessors. If the features that characterize the athletic outfit correlate with the extroversion assessment, then this message can be understood that athletic outfit triggers a certain reaction in a generic audience in terms of social signals. In order to individuate the attributes that most consistently originate social signals, deep learning technologies will come into play. One of the most attractive features of deep architectures is that they can be "opened" and "visualized", allowing to easily interpret what is codified into the internal layers [49], [25], [48], [32]. Exploiting these strategies, once annotations of

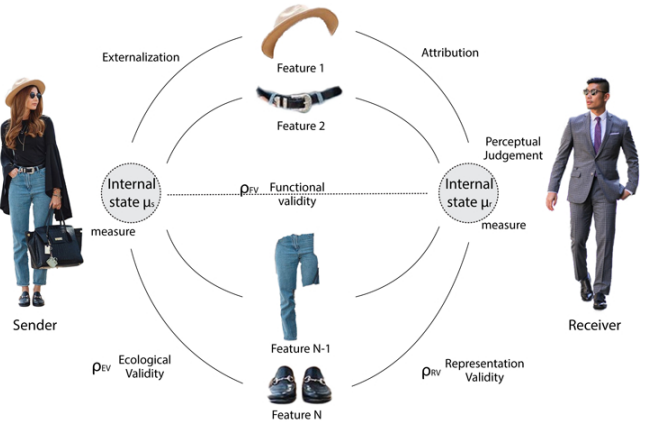


Fig. 3. The picture shows a simplified version of the Brunswick Lens Model adapted to the transmission of a social signal between a Sender and a Receiver.

social signals have been extracted from garment images, the goal would be that of feeding them into deep architectures, capturing the most discriminative visual patterns. In this way, the generic semantic label of "athletic outfit" can be explained in terms of behavioral cues (in the sense of [42]), like shape, color and texture attributes.

**A-Q3 - Is there an agreement between one's self-image and the impression conveyed to others through his/her clothing style?** It has long been known that *clothing affects how other people perceive us as well as how we think about ourselves*. This question asks whether there is a consistency between self-perception of an individual and perception of other people towards him/her.

**A-A3** - Often people choose what they wear as a means of self-expression. The individual measurement of the effect of clothing on self-perception and perception of others, has been studied previously by Heart [14]. However, the question of whether others perceive the desired message that the person wishes to communicate, has not been explicitly studied before. An experimental set up should first facilitate separate investigation of clothing effect over self-perception of an individual and perception of others over them and then study their correlation.

**A-Q4 - Which clothing behavioral cues are related to the social signal of the attractiveness?** This question asks if clothing style influences the perception of attractiveness of others. Attractiveness is the main social signal associated to physical appearance [42] and attractive people tend to be considered as having high status, good personality, and being extrovert.

**A-A4** - Clothing is considered as an indicator of socioeconomic status [39], [19] and personality traits [36]. According to Johnson et al. [18], the most investigated concepts using dress manipulations were *dress*, *status*, and *attractiveness* and listed the most experimented dress manipulation to study the effect of clothing on attractiveness concept as grooming, tidiness, makeup, and, natural physical appearance such as hair and eye color, height and, weight (see [18], Table 3). Clothing can be strongly related to arise perception of

attractiveness in people towards a person. To prove this hypothesis, a possible experiment would be that of capturing the influence of single clothing items, or multiple clothing elements arranged in an outfit towards the definition of an attractive person. In the same line with A-Q1, other behavioural cues should be considered, selected or masked, so to avoid explaining away effects. Also in this case, deep learning architectures and tools to visualize them ([49], [25], [48], [32], see A-A2) will help in segregate and explicitly individuate visual attributes that convey the impression of attractiveness.

**A-Q5 - How much impact clothing have on the other individuals first impression?** The importance of first impression comes more into attention in relation to its effect on the overall lasting impression. The last impression is what we will remember most about a social situation, however, one probably will not have a last impression if do not get the right first impression.

**A-A5 -** “You never get a second chance to make a first impression” as noted by Oscar Wilde. Although a large amount of cues aggregate together to form a first impression, we hypothesise clothing is a strong cue that eases the process for the people [2] to make a first impression. This quick judgment that happens in less than a minute [41], can lead us towards a set of assumptions about a set of personal traits of that person, such as attractiveness, likability, competence, and aggression [45]. Howlett et al. [17] studied the effect of clothing alone on the first impression and reported that clothing solely influences the first impression of the others even in limited exposure time. To detect the influence of individuals first impressions, the labeling of the *Exact Street2Shop* dataset (see A-A1) can be performed including the time dimension into play, enforcing the user to give a quick answer on the impression the clothing does trigger, explaining then by textual attributes the item(s) that leads him to such an answer.

### *B. Grounding clothing related social signals with scene context*

**B-Q1 - How much clothing-related cues help in capturing the context of a social interaction?** The idea here is to study how clothing items worn by people involved in a focused or unfocused interaction [34] can tell about the interaction itself.

**B-A1 -** People wearing outfits of a very similar kind, different from that of the rest of the crowd, are connected with a high chance, and this in turn helps in individuating the nature of a social interaction. Sport players with the same attire and supporters with the same t-shirts in a spectator crowd could be considered as an example of this connection. In this case, simple counting algorithms, specialized to finding similar items in a scene, may be of a great help [33]. However, the problem becomes more challenging when it comes to other types of interactions, namely ordinary exchanges in generic scenarios (waiting in a bus stop, attending a conference, etc.). An ideal solution is to develop models capable of first, assigning clothing visual

attributes to social scenarios (learning the most expected garments on the beach, in a Starbucks, during a conference etc.), and second, individuating similarities among outfits, from those very explicit (team outfits that are different only for the number depicted on the shoulders) to those more insidious to catch (individuating people that bring the same bag to individuate a social event). The interplay with social sciences lies in motivating the results from the learning stage of the model, that is, analyzing and interpreting the most emblematic garments for a particular interaction. Even in this case, deep learning technology will help, especially those architectures equipped with region proposal modules [30]. The novel idea here could be that of assuming the region proposal module as on-line evolving, drifting towards the detection of people exhibiting similar clothing.

**B-Q2 - How the clothing style drive people to socialize? Specifically, do people with the certain type of apparel socialize with similarly suited people?** This question asks whether our higher tendency to socialize with certain people is influenced by the clothing they wear, and if people tend to socialize more with similarly dressed people.

**B-A2 -** The approach towards answering this question is twofold. On one side, clothing in the same way as being considered as a flag to make visible a specific ideology, culture, or ethnicity, it also can be considered as a social catalysts among similarly dressed people. In an example, Nash [27] studied the influence of dressing on runners and stated that when two runners are dressed alike they engaged in an extended conversation as opposed to a short nonverbal greeting that occurred among runners that dressed differently from each other. On the other side, the effect of clothing on the people’s self-perception, leads to variations in their social relations. As an example, feeling comfortable is an important factor in a social interaction and clothing has the power to make a person feel comfortable or not. Simply, when clothing can be used as a criterion for judgment, people may unconsciously feel judged and act according to it. The connection with pattern recognition here lies in the approaches for detecting gatherings of people, which are proven to be very robust and versatile to diverse types of scenarios [34], [35]. Applying pattern recognition techniques for correlating clothing types of interactants will unveil possible affinities which may facilitate social interactions.

## V. CONCLUDING REMARKS

This work presents some research questions that are related to the investigation of the social signals associated to clothing. The outcome of these questions, other than filling a gap in the social signal processing literature, may have important relapses. For example, it could facilitate the design of online personal stylists able of indicating, on the one side, the type of impressions one’s outfit may trigger (associated to their particular body or posture), and on the other side, which are the most suitable outfits for attracting the attention of others. This has helpful applications in facilitating social interactions.

## REFERENCES

- [1] A. D. Adomaitis and K. K. Johnson. Casual versus formal uniforms: flight attendants' self-perceptions and perceived appraisals by others. *Clothing and Textiles Research Journal*, 23(2):88–101, 2005.
- [2] M. Barnard. *Fashion as communication*. Psychology Press, 2002.
- [3] G. V. Caprara, C. Barbaranelli, L. Borgogni, and M. Perugini. The big five questionnaire: A new questionnaire to assess the five factor model. *Personality and Individual Differences*, 15(3):281–288, 1993.
- [4] H. Chen, A. Gallagher, and B. Girod. Describing clothing by semantic attributes. In *European Conference on Computer Vision*, pages 609–623. Springer, 2012.
- [5] M. Cristani, A. Vinciarelli, C. Segalin, and A. Perina. Unveiling the multimedia unconscious: Implicit cognitive processes and multimedia content analysis. In *21st ACM International conference on Multimedia*, pages 213–222, 2013.
- [6] F. De Saussure. *The linguistic sign*. 1985.
- [7] W. Di, C. Wah, A. Bhardwaj, R. Piramuthu, and N. Sundaresan. Style finder: Fine-grained clothing style detection and retrieval. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 8–13, 2013.
- [8] J. Dong, Q. Chen, Z. Huang, J. Yang, and S. Yan. Parsing based on parselets: A unified deformable mixture model for human parsing. *IEEE transactions on pattern analysis and machine intelligence*, 38(1):88–101, 2016.
- [9] J. Dong, Q. Chen, X. Shen, J. Yang, and S. Yan. Towards unified human parsing and pose estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 843–850, 2014.
- [10] J. H. Fortenberry, J. MacLean, P. Morris, and M. O'Connell. Mode of dress as a perceptual cue to deference. *The Journal of Social Psychology*, 104(1):139–140, 1978.
- [11] C. Frith. Role of facial expressions in social interactions. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1535):3453–3458, 2009.
- [12] M. Hadi Kiapour, X. Han, S. Lazebnik, A. C. Berg, and T. L. Berg. Where to buy it: Matching street clothing photos in online shops. In *IEEE International Conference on Computer Vision*, pages 3343–3351, 2015.
- [13] B. Hannover and U. Kühnen. The clothing makes the self via knowledge activation. *Journal of Applied Social Psychology*, 32(12):2513–2525, 2002.
- [14] A. Heart. *Apparel and its impact on self-perception*. 2011.
- [15] R. Hodge and G. Kress. *Social semiotics*. Polity Press, 1988.
- [16] K. Hogan. The secret language of business. *The Secret Language of Business: How to Read Anyone in 3 Seconds or Less*, pages 1–16.
- [17] N. Howlett, K. Pine, I. Orakcioglu, and B. Fletcher. The influence of clothing on first impressions: Rapid and positive responses to minor changes in male attire. *Journal of Fashion Marketing and Management: An International Journal*, 17(1):38–48, 2013.
- [18] K. K. Johnson, J.-J. Yoo, M. Kim, and S. J. Lennon. Dress and human behavior a review and critique. *Clothing and Textiles Research Journal*, 26(1):3–22, 2008.
- [19] J. Kerr. Dress as a status symbol. *African American dress and adornment*, pages 93–103, 1990.
- [20] M. H. Kiapour, K. Yamaguchi, A. C. Berg, and T. L. Berg. Hipster wars: Discovering elements of fashion styles. In *European conference on computer vision*, pages 472–488. Springer, 2014.
- [21] P. Kohli, P. H. Torr, et al. Robust higher order potentials for enforcing label consistency. *International Journal of Computer Vision*, 82(3):302–324, 2009.
- [22] X. Liang, S. Liu, X. Shen, J. Yang, L. Liu, J. Dong, L. Lin, and S. Yan. Deep human parsing with active template regression. *IEEE transactions on pattern analysis and machine intelligence*, 37(12):2402–2414, 2015.
- [23] S. Liu, J. Feng, C. Domokos, H. Xu, J. Huang, Z. Hu, and S. Yan. Fashion parsing with weak color-category labels. *IEEE Transactions on Multimedia*, 16(1):253–265, 2014.
- [24] S. Liu, X. Liang, L. Liu, X. Shen, J. Yang, C. Xu, L. Lin, X. Cao, and S. Yan. Matching-cnn meets knn: Quasi-parametric human parsing. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1419–1427, 2015.
- [25] A. Mahendran and A. Vedaldi. Understanding deep image representations by inverting them. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 5188–5196, 2015.
- [26] A. C. Murillo, I. S. Kwak, L. Bourdev, D. Kriegman, and S. Belongie. Urban tribes: Analyzing group photos from a social perspective. In *Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 28–35, 2012.
- [27] J. E. Nash. Decoding the runner's wardrobe. *Conformity and Conflict*, eds. James P. Spradley and David W. McCurdy, Boston: Little, Brown, 172:185, 1977.
- [28] Y. S. M. Owyong. Clothing semiotics and the social construction of power relations. *Social Semiotics*, 19(2):191–211, 2009.
- [29] W. T. Pauline. *The victorian era fashion history*, 2006.
- [30] S. Ren, K. He, R. Girshick, and J. Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems*, pages 91–99, 2015.
- [31] R. B. Rubinstein. *Dress codes*. Boulder, CO: Westview, 1995.
- [32] W. Samek, A. Binder, G. Montavon, S. Lapuschkin, and K. R. Müller. Evaluating the visualization of what a deep neural network has learned. *IEEE Transactions on Neural Networks and Learning Systems*, PP(99):1–14, 2016.
- [33] F. Setti, D. Conigliaro, M. Tobanelli, and M. Cristani. Count on me: learning to count on a single image. *IEEE Transactions on Circuits and Systems for Video Technology*, PP(99):1–1, 2017.
- [34] F. Setti, C. Russell, C. Bassetti, and M. Cristani. F-formation detection: Individuating free-standing conversational groups in images. *PloS one*, 10(5):e0123783, 2015.
- [35] J. Shao, C. Change Loy, and X. Wang. Scene-independent group profiling in crowd. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2219–2226, 2014.
- [36] M. L. Slepian, S. N. Ferber, J. M. Gold, and A. M. Rutchick. The cognitive consequences of formal clothing. *Social Psychological and Personality Science*, 6(6):661–668, 2015.
- [37] G. B. Sproles and L. D. Burns. *Changing appearances: Understanding dress in contemporary society*. Fairchild Publications, 1994.
- [38] E. Stephan, N. Liberman, and Y. Trope. The effects of time perspective and level of construal on social distance. *Journal of Experimental Social Psychology*, 47(2):397–402, 2011.
- [39] R. Sybers and M. E. Roach. Sociological-research-clothing and human-behavior. *Journal of Home Economics*, 54(3):184–187, 1962.
- [40] A. R. Timming and D. Perrett. Trust and mixed signals: A study of religion, tattoos and cognitive dissonance. *Personality and Individual Differences*, 97:234–238, 2016.
- [41] A. Todorov, M. Pakrashi, and N. N. Oosterhof. Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition*, 27(6):813–833, 2009.
- [42] A. Vinciarelli, M. Pantic, and H. Bourlard. Social signal processing: Survey of an emerging domain. *Image and Vision Computing*, 27(12):1743–1759, 2009.
- [43] A. Vinciarelli and A. S. Pentland. New social signals in a new interaction world: the next frontier for social signal processing. *IEEE Systems, Man, and Cybernetics Magazine*, 1(2):10–17, 2015.
- [44] S. Vittayakorn, K. Yamaguchi, A. C. Berg, and T. L. Berg. Runway to realway: Visual analysis of fashion. In *Winter Conference on Applications of Computer Vision*, pages 951–958. IEEE, 2015.
- [45] J. Willis and A. Todorov. First impressions making up your mind after a 100-ms exposure to a face. *Psychological science*, 17(7):592–598, 2006.
- [46] K. Yamaguchi, M. H. Kiapour, L. E. Ortiz, and T. L. Berg. Retrieving similar styles to parse clothing. *IEEE transactions on pattern analysis and machine intelligence*, 37(5):1028–1040, 2015.
- [47] Y. Yang and D. Ramanan. Articulated pose estimation with flexible mixtures-of-parts. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1385–1392, 2011.
- [48] J. Yosinski, J. Clune, A. Nguyen, T. Fuchs, and H. Lipson. Understanding neural networks through deep visualization. *arXiv preprint arXiv:1506.06579*, 2015.
- [49] M. D. Zeiler and R. Fergus. Visualizing and understanding convolutional networks. In *European conference on computer vision*, pages 818–833. Springer, 2014.