Gestalt Processing in Human-Robot Interaction: A Novel Account for Autism Research

Maya Dimitrova
Institute of Systems Engineering and Robotics, Bulgarian Academy of Sciences, Department "Hybrid Systems", Bl.2, Acad. G. Bonchev Str., POB 79, 1113 Sofia, Bulgaria
dimitrova@iser.bas.bg

Abstract

The paper presents a novel analysis focused on showing that education is possible through robotic enhancement of the Gestalt processing in children with autism, which is not comparable to alternative educational methods such as demonstration and instruction provided solely by human tutors. The paper underlines the conceptualization of cognitive processing of holistic representations traditionally named in psychology as Gestalt structures, emerging in the process of human-robot interaction in educational settings. Two cognitive processes are proposed in the present study - bounding and unfolding - and their role in Gestalt emergence is outlined. The proposed theoretical approach explains novel findings of autistic perception and gives guidelines for design of robot-assistants to the rehabilitation process.

Keywords: Human-robot interaction, Gestalt processing, perception, autism, cognitive technology, robot-assistants, rehabilitation robotics

1. Introduction

With the present-day increased interest in robotic systems entering rehabilitation – physical and psychological – it is necessary to formulate novel principles of communication between the user and the robot in settings, where the robot has become an assistant to the entire rehabilitation process (Bailey-Van Kuren, 2007). These principles are cognitive in nature and help understand in a contemporary manner the cognitive impacts of novel interactive technology such as touchscreen-based user-robot interaction (Hashimoto et al., 2011), robots understanding voice (Kim et al., 2009) and gestures (Hafner et al., 2005), implemented and used in clinical settings.

The proposed theoretical framework of the cognitive processing underlying present-day human-robot interaction is based on the following theoretical definition of the Gestalt proposed by Barry Smith: "A Gestalt is a whole of relations, but in certain circumstances only part of this whole may be perceived – and this part may be a Gestalt in its own right. Indeed, it is only in certain simple cases, for example simple visual patterns that the entire Gestalt can be perceived in one intuitive glance. In more complicated cases this is not possible, and the greater the manifold of relations between the parts of a given field the less it is possible to grasp all relations simultaneously. We can grasp the Gestalt only if we are somehow able, by a cumulative process involving the operations of memory, to unify everything in one intellectual glance, and a discursive process of this sort is indeed indispensable if we are to grasp a melody or other Gestalt involving any sort of temporal succession" (Smith, 1988, p. 25).

Human cognition – from a developmental perspective – accumulates the incoming information into Gestalts – entities representing the whole of it into *one intellectual glance* in recursive and adaptive manner. With autism, for example, the disruption of the smooth flow of this recursive process is the main educational obstacle, and it is where robots have been extremely useful in coping with repetitive actions and attention shifts ((Bailey-Van Kuren, 2007; Huskens et al., 2013; Dimitrova et al, 2012; Robins et al., 2006). Being complex technological devices with many details capturing attention yet being stable and predictable – makes robots ideal for providing guidance in education and in practicing novel cognitive and social skills.

The current paper focuses on the principles as well as on newly discovered elements of Gestalt phenomena in their relevance to rehabilitation robotics. The cognitive processes, assumed

Gestalt in their phenomenological nature, emerge in the context of the psychological reality of the immediate cognition, implicit (spontaneous) retrieval from memory, and the autobiographical nature

of the human-agent/robot interaction (Dimitrova &Wagatsuma, 2011). Two cognitive processes are proposed in the present study - called (internal) *bounding* and *unfolding*. The 'unfolding' term is first used by C. von Ehrenfels (1988, p. 82) to explain the unfolding of a Gestalt in time. Here the word *unfolding* is used to designate an internal dynamic process, operationally defined and described in the present investigation of the emergence of the cognitive 'whole, being more than the sum of its parts', as Gestalt is being frequently defined.

Often the term Gestalt is used interchangeably with the term "emergent whole" (Johansson, 1998). The emergence of a cognitive Gestalt structure adds dynamical and psychophysical forces, which are different from the static notion of the "emergent whole". An eminent example for the dynamic nature of the emergent process is the Necker cube, which cannot be perceived as static, but rotates in front of our eyes to the complete exhaustion of the eye gazing process (figure 1).

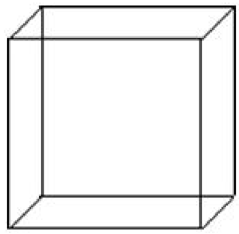

Figure 1. Necker cube depth illusion (Adapted from [http://en.wikipedia.org/wiki/Necker_cube])

The ambiguous figure interchangeably displays its front or back wall and this constant process is the spontaneous reaction of the human visual system. None less eminent is the Zeigarnik phenomenon (1927), demonstrating that people have superior memories for the interrupted cognitive activities, than for the completed ones (Zeigarnik, 1988, p. 300). The memory for the unresolved task stays longer, being constantly retrieved in its various aspects by some internal to the 'task figure' cognitive force, which is an energy consuming process like looking at the Necker cube.

In the present paper the analysis focuses on showing that education is possible through robotic enhancement of the Gestalt processing in children with autism, which is not comparable to alternative educational methods such as demonstration and instruction provided solely by human tutors – and therefore much more productive and useful in education of children with special needs.

2. Robotic Technology in Autism Research – Enhancing the Composition of Gestalt Structures in the Process of Cognition

Recently a number of theoretical and experimental paradigms have been proposed, stating that educational rehabilitation – in particular for autistic spectrum conditions (ASC) – is much more enhanced if mediated by complex cognitive-technology facilities like robots (Huskens et al., 2013; Robins et al., 2006; LeGoff, 2004; Perlin, 2004; Tentori & Hayes, 2010) than by the traditional teacher-student-parent educational paradigm. The assumption, presented here, is that socially-mediated learning is based on deep cognitive Gestalt structures, maintaining the basic relations to be understood through the explanations elicited from the teacher. But because social interaction is emotionally demanding in its influence on the student, children with autism will better socialize if the process is mediated by a robot-assistant, being less emotionally redundant and more cognitively focused.

LeGoff developed a measurement tool to assess social interaction in autistic children during therapy (LeGoff, 2004). In his research on using Lego© as a therapeutic medium, social interaction is distinguished in three elements: "(1) initiation of social contact with peers, reflective of social interest and motivation for social contact; (2) duration of social interaction, which reflects the development of communication and play skills; and (3) decreases in autistic aloofness and rigidity, with development of age-appropriate social and play behaviors" (LeGoff, 2004, p. 557). He proposed an indicator of social interaction, called a "self-initiated social contact" (SISC) measure. SISC can be observed through a semi-transparent mirror, or by sampling video recording into short episodes. Legoff applied the methodology to observation of children with ASC and reported increased amount of self-initiated social contact among these children while playing with Lego©. This has subsequently become a popular methodology for quantitative evaluation of the educational influence of technology to children with difficulties in social communication.

Ken Perlin is the founder of Games for Learning Institute employing many computer simulations to education. Among numerous simulations, one deserves special attention for implementing a cognitive theory of facial expression of emotions – the theory of Paul Ekman, who proposed a Facial Action Coding System providing a list of functional descriptions of expressions of the human face (Ekman et al, 2002). The demo of a girl's face can represent the following emotional expressions: frightened; kiss; disappointed; sleeping; annoyed; surprised; happy; arrogant; angry and talk. The interface can manipulate the movement of the following face elements: brows, lids, gaze, head and mouth. These can be manipulated in different directions or with commands like "open", "wide" and "sneer". The Facedemo is used by the Do2Learn Project funded by the US National Institutes of Health Small Business Innovative Research Grant. This applet has helped autistic children learn the association between a holistic facial expression of emotion (configural processing) and the individual parts of the face signifying emotional expressions (featural processing) – like raised eye-brows in surprise – and manage to successfully integrate them in understanding behavioral patterns (http://www.mrl.nyu.edu/~perlin/facedemo/). Although the applet is two-dimensional in nature – three-dimensionality is achieved by using the laws of perspective and the experience is of immersion in the created reality. It is both entertaining and educative and useful for describing facial expressions of people's emotions and enhancing the social competence of the learner (Perlin, 1997).

MOSOCO is an emergent technology implemented in a smartphone called "Mobile Social Compass" (Escobedo et al., 2012). Six basic social skills are being encouraged by prompting the user to initiate social contact. The menu displays symbols for the basic social skills – eye contact, space and proximity, start an interaction, asking questions, sharing interests and finish an interaction. The MOSOCO application has turned out to be an extremely useful tool as an online prompt in starting, maintaining and finishing social interaction for both typical and autistic students, as well as to anyone that feels need for improving their social competence.

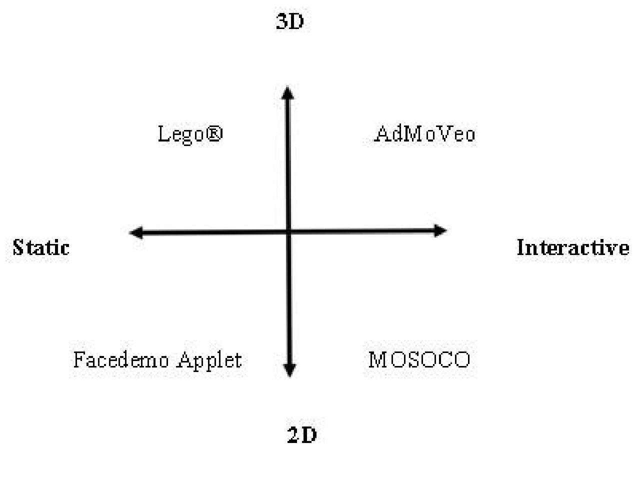

Figure 2. Taxonomy of the educational technologies for children with ASC

A taxonomy of the existing cognitive technologies based on the presence of two design factors – static vs. interactive design and two vs. three-dimensional design is presented in figure 2. The existing technological tools can serve the educational needs in general and the needs of the autistic children in particular. Figure 2 depicts four examples of technological tools extensively used to teach social skills to children with ASC – Lego®, the Facedemo Applet designed by Ken Perlin, the mobile platform for instructing children MOSOCO and a multi-agent platform of interactive robots for teaching social skills AdMoVeo from www.admoveo.nl (counterclockwise in figure 2). They are chosen as examples of the role of cognitive technology in improving different social skills integrated in the general educational level of children – self-initiated social contact, description of facial expressions and the respective emotions they convey, provision of socially significant cues in the process of communication and interactive robots instructing and rewarding collaborative games of children with ASC. These 4 abilities comprise complex and multi-level social cognition ability (or Gestalt) based on the involved sophisticated technology. The robotic demonstration and instruction is a valuable enhancement of the Gestalt processing in children with ASC aiming at involving the use of the mirror neuron system in the brain (Rizzolatti & Fabbri-Destro, 2010) which is spared in autism, yet atypically modulated (Haffey et al., 2013) to provide better epigenetic adaptation in the process of child development (Tramacere et al., 2013).

3. Cognitive Processing Involved in the Gestalt Dynamics

The emergence of holistic entities like images from patterns of pixels and color dots was explained in (Johansson, 2006). The elements of the base level are called "concrete collections", being organized according to some rule, whereas the 'image' that is emergent form these concrete collections is organized at the same time according to some ontologically supervenient rule. Johansson illustrates the idea of the emergent whole with the example of the popular smiley :-). When the elements – the punctuation marks - are placed in arbitrary order, there is a basic rule to preserve their individual format and their spatial relations. This rule may hold for their spatial configuration and it can define a configuration, independent from the one which produces the idea of a face in the viewers' perception. The face configuration is dependent on the individual configurations and formats of the elements, but emerges on a different – supervenient – level, obeying a different rule on this next level – a rule describing a face. What is important is to bind inferentially the two levels and to preserve the deeper level - the meaning. The idea of a face is thus bounded by the format of the individual elements and not that much by their spatial relations any more, since the supervenient rule requires this. As it is formulated by Barry Smith: "That is, just as we conceive the complex whole in [A] as possessing a certain characteristic whole-property, so we can conceive the different parts of this complex as possessing their own characteristic part-properties in virtue of which they come to make up that total whole which is the original Gestalt" (Smith, 1988, p. 54).

To illustrate the preserved part-properties in the *smiley* Gestalt, just imagine a smiley on a graffiti wall or even a smiley in a graffiti paining (similar to the figure-ground effect). Color, paint and texture have their own complex part-properties and nevertheless belong to the smiley as a face. From a Gestalt perspective it is necessary to mention another aspect of the smiley example of an emergent whole – the *simplicity* notion of Max Wertheimer (1958). Nowadays a smiley is a two-element entity – happy :) and sad :(face – with maximally shrunken base and richness of meaning at the supervenient level – expressing simultaneously a face and a palette of emotions – and is a Gestalt.

The smiley Gestalt is the result from a dynamic (evolved in time) cognitive process, it is immediately given in cognition, memorable, emotionally rich and socially relevant and reflects the special kind of Gestalt *complexity* as defined by Edwin Rausch (1988). Conventional representations of holistic entities like smileys or *novel robotic faces* come to life because they capture essential Gestalt qualities of the perceived image. For example, in figure 3 the robotic faces resemble the smileys in terms of the evoked internal/emotional reactions.

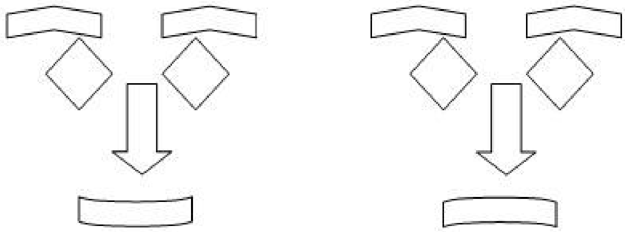

Figure 3. Robotic faces, similar to smiley emoticons

In a study on recognition of emotions, expressed by human vs. robotic faces, evoked potentials were obtained from the visual cortex of people (Dubal et al., 2011). The study was the first registration of brain response to emotion, but not to type of face – machine or human. This posed a problem to theories, stating that social recognition comes after perceptual recognition. It is more feasible to adopt a social motivation account of the phenomenon like in (Chevallier et al., 2012) and propose the emergence of an emotional Gestalt in parallel with the perceptual one. Human and machine faces with neutral expressions were not distinguishable in terms of the magnitude of the evoked potentials, whereas the emotional expressions were distinguishable from the neutral faces – either human or machine – in typical experimental subjects. This is an important finding to relate to face recognition in people with ASC (Behrmann et al., 2006).

The forwarded in the present paper hypothesis is that the process of constant translation from *emotion recognition* to *face recognition* and vice versa is disrupted in the face Gestalt, in a manner similar to as if the perceptual motion of the Necker cube is being disrupted. This account supports the hypothesis that autistic perception is a matter of cognitive style rather than cognitive deficit, meaning that autistic features are distributed in the general population of people, but not concentrated in the syndrome only (Happe, 1999). Therefore, cognitive support to this translation process can improve the cognitive ability of the learner.

4. Simulataneous Configural and Featural Processing of the Face Gestalt

In 1980 Peter Thompson proposed a new experimental paradigm for investigation of perception, called "face thatcherization" (also named "Thomson illusion") (Thompson, 1980). Imagine that the following face, depicted in figure 4, is a photo of the then UK Prime Minister Margaret Thatcher.

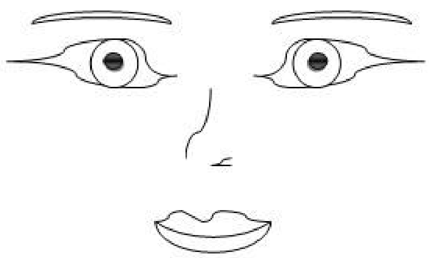

Figure 4. Main elements of a face, according to the feature-based processing theories

According to the feature-based processing theories of human faces the main elements, noticed and remembered in a face, are the eyes, the nose and the mouth (Thompson, 1980; Anstis, 2005a). If, however, we distort some of the elements of a face, these should influence perception, regardless of the position of the image – upright or reversed – from the observer viewpoint. Figure 5 presents the reversed image of the face on the left and the reversed distorted face on the right. The distortion was achieved by rotating the eyes of the image in the vertical direction.

Figure 5. The distortion is barely noticeable if the faces are viewed in the reversed position

Quite surprisingly, if the distortion is viewed in the normal upward position, it evokes strong emotional response to the distorted face to the right in figure 6.

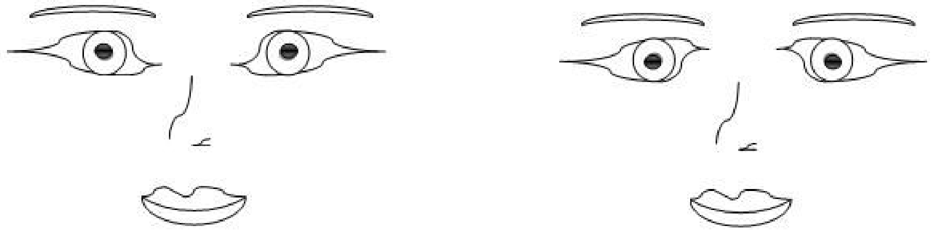

Figure 6. Noticeable subjective response to the distorted face to the right

This incongruent perception of upright and reversed distorted faces was interpreted by the author of the illusion as favoring the configural face-processing account – i.e. that distortions become noticeable on a second view, whereas dominant is the perception of the entire configuration as in viewing reversed images of faces. With realistic photos the emotional reaction is much stronger, which made the author remark ironically at the end of the paper: "Further research into this illusion might help determine whether face recognition is a serial or a parallel process, is achieved by face specific analysers, searching for distinguishing characteristics or by some perceptual gestalt. It might even tell us something about Margaret Thatcher" (Thompson, 1980, p. 483).

Stuart Anstis designed the so called "Blair illusion" as a révérence to Peter Thompson (Anstis, 2005a). It is an example of the same phenomenon as "face thacherisation" in the visual contrast domain. In the contrast domain – a photo negative of an image is used. The features – eyes and mouth – are contrasted to the rest of the image. When the image is viewed in its reversed to normal contrast, the non-reversed elements are not perceived as grotesque (corresponding to viewing the inverted image of a thatcherised face). Quite on the contrary - and in full parallel with the Thatcher effect – when the image is viewed in its normal contrast, with inverted eye and mouth contrast, the effect is equally grotesque. In 2005 Stuart Anstis concluded his paper on the discovery of an effect in the visual contrast domain mirroring the Thatcher effect in the positional domain, which he called Blair effect, with the following sentences: "Lewis and Johnston (1997) found that the Thatcher illusion occurs, in albeit a reduced form, when photographic negatives are used. This would suggest that there are similarities between the Blair illusion and the Thatcher illusion (in much the same way as there are similarities between Blair and Thatcher)" evidently referring to social Gestalt phenomena outside laboratory research.

A study on Thompson illusion is presented in (Milivojevic et al., 2003). In this study the effects of rotation of a "thatcherised" face is measured by behavioral tests and event-related potentials. The main results of the study are the following. The authors compared an objective measure of the effect of rotation with its subjective counterpart. The objective measure was the report of the participants how bizarre the rotated image seemed to them. It has turned out that bizarreness depends on the relative orientation – and is gradual, even though not perfectly matching the angle of rotation. The subjective measure was to report the degree of unpleasantness of the face. As a result, the emotional reaction was abrupt rather than continuous. The study is important in outlining the simultaneity of the objective and subjective processing of faces by our cognitive system. Defining bizarreness of an image is a process of decomposition and reality check of the elements of this image. Defining pleasantness of the image is *unfolding* of the idea of a pleasant face and folding of the bizarre image into a configural entity corresponding to an approximately pleasant or unpleasant face. This experiment empirically captures an important Gestalt principle of co-coinciding with the reality and co-continuing with it (Smith, 1988). Reality is commonly analysed by decomposing the *givenness* of the object – a face – into elements – to be further analysed and further decomposed. The Gestalt *unfolding* process coincides with the decomposition process by extracting the approximate central parts of the objective configuration – the ones produced by the decomposition – to fold them in a cognitive entity without losing the meaning of this entity – the pleasant face or the unpleasant face.

A special case of Gestalt processing is the perceiving of illusions. Illusions make us see things or processes that are not there – for example the Kanizsa square like the one depicted in figure 7.

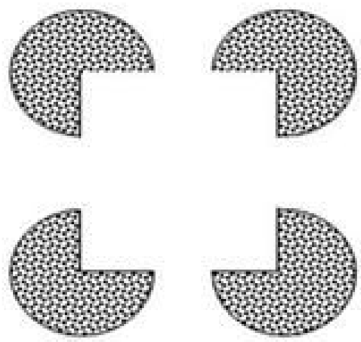

Figure 7. Kanizsa square makes us see a non-existing figure – white square (Adapted from [http://en.wikipedia.org/wiki/Optical_illusion]

The Thatcher illusion and the Blair illusion are not, in a sense, perceptual *illusions* per se. They are not visual illusions, but they are, rather, perceptual effects, which – in the context of the initial persons' photographs – are not even perceptual, but emotional in the evoked response. As explained above – human perception is sensitive to the position of an image – the upright and the reversed position – of a human face. It is also sensitive to the corresponding position of the elements of the image – the eyes and the mouth. The compared effects of viewing different orientations of the whole image together with the inverted elements – also reveal parallel processes of visual scanning for the whole configuration on the one hand and for the individual elements – on the other. The complexity of simultaneously scanning for two things at a time is revealed by the non-homogeneity of these processes in time and space as measured by the response reaction times, *bizarreness assessment* (objective report) and *pleasantness judgment* (subjective report).

Experimental studies on cell response of the monkey's brain to artificial faces have revealed that areas of the brain responding to luminance contrast, shape of the face, and visual attention are linked (Kandel, 2013). These findings support the Gestalt understanding of the complexity of face recognition (shape) related to the Kanizsa illusion perception (luminance) and attention (featural processing). In autism the constant cognitive movement from one visual facet of the image to another is slower and more serial. People with ASC do not perceive the impossible figures as such (Liu et al., 2011). They are superior in finding embedded figures in an image (Bolte et al., 2007) and less emotionally involved in the process of face recognition (White et al., 2009). The general hypothesis is that the bizarreness assessment would be confused with pleasantness judgment in autism where these processes are performed serially and need a special prompt.

5. Orientation of the Visual and Lexical Gestalts – Novel Cues

In an experimental study on visual perception, addressing directly Gestalt phenomena, a new Gestalt cue for figure-ground assignment was introduced (Vechera et al., 2002). The foreground versus the background organization is a strong determinant for decisions on objects seen among image elements. A well-known set of perceptual cues that are often called Gestalt cues are the size or area, the symmetry and the convexity vs. concavity judgments. It is generally assumed that figures are 'small, symmetrical and convex'. The authors asked the question whether these cues are all that are necessary for a region of the image to be judged as a figure. The main result of this study is that regions in the lower portion of a stimulus array appear more figure-like than regions in the upper portion of the display.

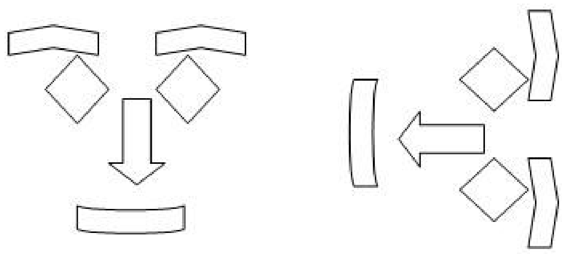

Figure 8. Machine faces, perceived as more figure-like(left) and less figure-like (right)

The cognitive effect is that an ambiguous image more often is judged as containing a figure when positioned up-down, rather than when positioned left-right. Along the results of their experiments, the machine (robotic) face to the left in figure 8 should be perceived as more figure-like, than the one to the right. The authors conducted 8 experiments and found that this was a robust phenomenon. The effect is relative to the horizontal horizon line and the authors suggest that the lower region preference is linked to the pictorial depth perception cues. The authors also discuss the possibility that neural mechanisms in the visual cortex may exist to account for both depth perception and lower region figural preference (Vechera et al., 2002).

In the lexical domain a similar effect of holistic word processing is described in (Anstis, 2005b). The viewers were presented with pairs of three-letter words in quick succession and asked to report if the upper halves of the successively presented words were identical. Surprisingly, even when the upper halves of the words were orthographically identical, the error rate was reliably higher than expected and in comparison with matching identical successive words. As the author of the study Stuart Anstis points out: "students were processing the words not as separable parts, but holistically as perceptual units that could not be perceptually split apart. These results show that in normal circumstances, the visual system cannot, or does not, divide words into upper and lower halves" (Anstis, 2005b, p. 239).The author relates the results of his study to studies of visual perception of faces as evidence that the mechanism of holistic processing in the visual and the lexical domains is essentially the same.

Figure 9. The impossible figure (right) is not noticeable as such at first glance

This regularity of cognition is about the geometry of the internal representation of the external world, which somehow most often is approximating the Euclidean space according to L.M. Vekker (1976). In his theory of the Gestalt in perception and thinking, Vekker postulated that sensations, perceptions and imagination are in a constant process of *unfolding* to match and take the exact places of the physical objects in the external environment. The geometry of the internal representation of the world is abstract and experiential at the same time. The experiential aspect of the process of immediate cognition can be seen by viewing figure 9. Figure 9 right is, at first glance, a quite convincing example of stairs similar to the stairs on the left. Even the context (the cat on the ceiling) is noticed on a second scan. The inverted stairs are more likely to remind us of Escher's works than, for example, the normally positioned ones. This effect supports the idea of the semantic depth of spatial learning, relating the experience of art to perceiving of Gestalts.

See http://en.wikipedia.org/wiki/M._C._Escher

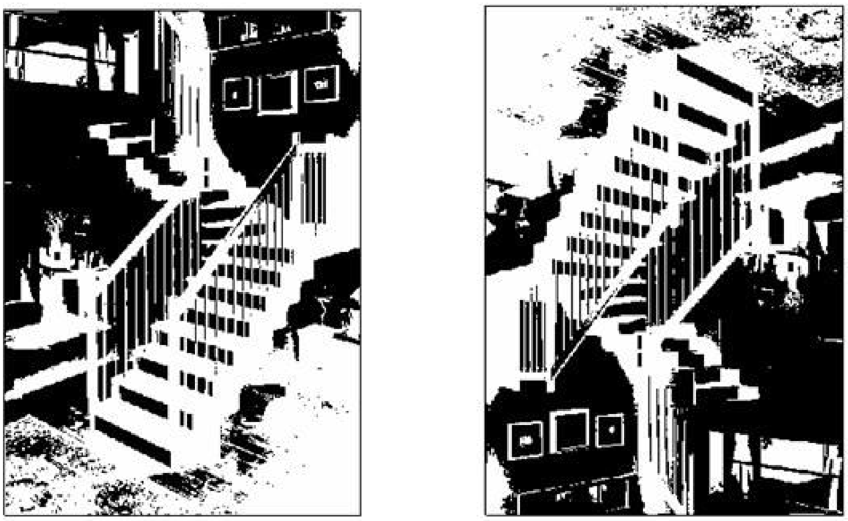

Figure 10. Otsu segmentation of the real (left) and the impossible (inverted) figure (right)

A possible computational interpretation of this effect of dominance of the configural processing over feature-based processing in typical vision can be provided by the Otsu method of image segmentation as illustrated in figure 10 (Otsu, 1979). It is obtained with Scilab's image processing toolbox (Galda, 2011) and illustrates the dominance of the vertical over horizontal processing in the human visual system. The Otsu method assumes that a gray level image can be divided into foreground and background as a binary (black-white) image by minimizing the intra-class variance defined as a weighted sum of the variances of the two classes. The transformed images from figure 9 into Otsu images in figure 10 produce dominant image configurations, where the details (like the cat on the ceiling) are lost. People with ASC are superior in local processing at the expense of global processing and are, therefore, less susceptible to Gestalt grouping (Brosnan et. al, 2004). It seems that the Otsu-like computation in the brain is less efficient in people with ASC than in the typical cognitive processing of human vision.

6. Cognitive Processing of Social-Communication Signals

A possible cognitive architecture and formalization of the process of learning via multisensory integration is presented in figure 11. The formal description of the proposed cognitive architecture, capable of interpreting social-communication signals, signs and symbols, is based on multisensory integration at the level of perception, parallel processing at the level of interpretation and decision making followed by verbalization, as well as performing an action (eye contact, gesture, mimicking) at the level of behaviour.

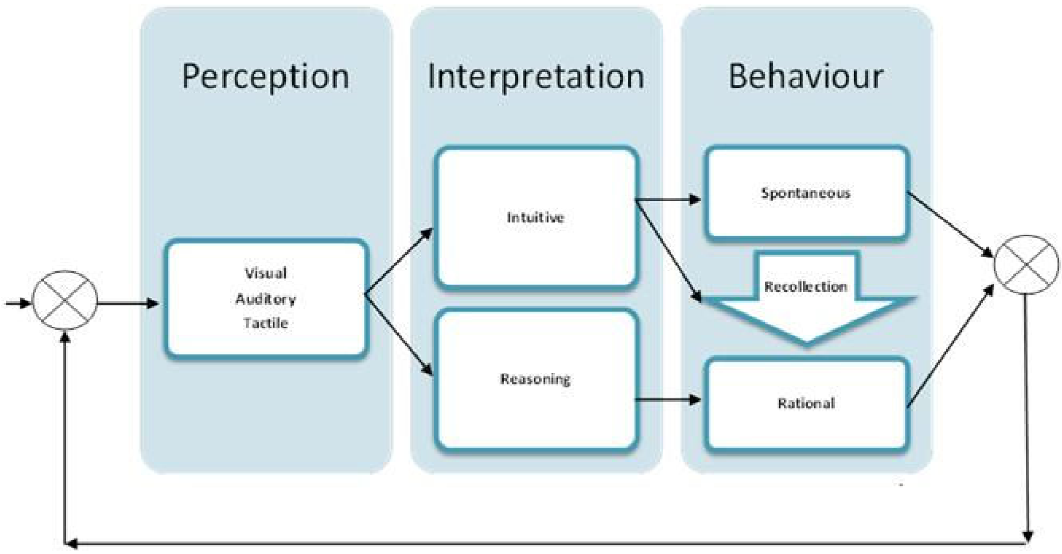

Figure 11. Cognitive architecture of the process of social signals perception

The cognitive architecture evolves in the closed loop of the sensor-effector continuum,

$$Dev = Cog + Beh \quad (1)$$

where *Dev* is the current stage of learning on the part of the cognitive architecture.

$$Cog = P(v_i,a_i,t_k) + \gamma[I(i_i) + I(r_j)] \quad (2)$$

where *Cog* is the combined multisensory perception $P(v_i$, - visual, a_i–auditory и t_k – tactile) plus interpretation of the incoming information *I*, modulated by a subjective index γ (r_j is feedback correction).

$$Beh = \tau[B(s_i) + B(s_i)' + B(r_j)] \quad (3)$$

where *Beh* is the combined behavioural response *B*, modulated by a subjective index τ and B' is a spontaneously recollected concept, otherwise 0.

The different degrees of preservation of the cognitive dynamics will lead to different predictions about the pace of learning and the effectiveness of the education in robot-tutor or human-tutor situations. As a result educational methods can be developed, adapted to the specific cognitive needs of the child.

The interplay of the processing in the different cognitive blocks produces the Gestalt phenomena of dynamically folding and unfolding the elements of an image into a whole and at the same time processing the individual elements in a holographic manner. Disruption of this dynamic process leads to effects of inadequately assigned cognitive resource. In (Dimitrova et. al, 2012) the limited resource hypothesis is forwarded assuming that the attention span is narrow yet deep in children with ASC. They are less flexible in shifting attention, which is supported by the conducted experiment.

In (Feil-Seifer & Mataric, 2012) a robot-assistant capable of automatic classification of positive vs. averse child behavior and socially-aware navigation is implemented. Robots of this kind can be augmented with more complex understanding of the emergent learning Gestalts on the one hand, and the process of evolving of the learning in autism – on the other.

In (Huskens et al., 2013) direct comparison between robot intervention and human-tutor intervention was made in order to try and assess the effectiveness of using a robot as an educational method. The amount of self-initiated questions increased and was maintained in both groups – control (human-tutor) and experimental (NAO robot), which led to the conclusion that both interventions were effective. In (Duquette et al, 2008) some of the interactions like eye contact and physical proximity were more effective in the robot-tutor condition. This is a promising result in support of the present claim that enhancing the Gestalt processing would produce better results in education than the traditional approach of presenting novel knowledge element by element. Holistic processing in parallel with processing of individual elements is a plausible way of complex natural computing performed by the human brain.

4. Conclusion

The presented comparative analysis of formation of cognitive representations of the environment – social and physical – as aspects of the Gestalt understanding of the surrounding world in typical and people with ACS – reveals that the complexity of the adaptation to the world can be usefully mediated by complex cognitive technologies like robot-assistants to the rehabilitation process. The proposed theoretical approach explains novel investigations in the Gestalt phenomenon as well as findings of perception in ASC that are helpful for guidelines for design of robot-assistants to the educational process. With the inclusion of robots in the educational and instructional activities of all possible aspects of life in the near future, and with the increased cognitive competence implemented in technology, people of different learning styles and educational needs will rely on personalized and individualized instruction. Further investigation and deeper understanding of the Gestalt phenomena in human cognition will be very useful for building truly intelligent cognitive technologies of the future.

Acknowledgement: I am grateful to Professor Barry Smith and Professor Ingvar Johansson form IFOMIS, Saarland University, Germany, for sharing their work and ideas on the Gestalt theory. This work was partially supported by EEA Grant No D03-90/27.05.2015 of BG09 Programme of the Financial Mechanism of EES 2009-2014, Measure "Projects for Interinstitutional Cooperation".

References

Anstis, S. (2005a). Homage to Peter Thompson: The Tony Blair illusion. *Perception*, 34, 1417–1420.

Anstis, S. (2005b). Holistic word processing. *Perception*, 34, 237–240.

Bailey-Van Kuren, M. (2007). Robotic solutions in pediatric rehabilitation. In S. S. Kommu (Ed.), *Rehabilitation Robotics* (pp. 1-12). I-Tech Education and Publishing, Vienna, Austria.

Behrmann, M., Thomas, C., & Humphreys, K. (2006). Seeing differently: Visual processing in autism. *Trends in Cognitive Sciences*, 10(6), 258-264.

Bolte, S., Holtmann, M., Poustka, F. Scheurich, A., & Schmidt, L. (2007). Gestalt perception and local-global processing in high-functioning autism. *Journal of Autism and Developmental Disorders*, 37, 1493–1504.

Brosnan, M., Scott, F.J., Fox, S., & Pye, J. (2004). Gestalt processing in autism: failure to processperceptual relationships and the implications for contextual understanding. *Journal of Child Psychology and Psychiatry*, 45 (3), 459–469.

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E.S., & Schultz, R.T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, 16(4) 231-239.

Dimitrova, M. &Wagatsuma, H. (2011). Web agent design based on computational memory and brain research. In: N. Tang (Ed.) *Information Extraction from the Internet* (pp. 35-56), Hong Kong, China, iConcept Press Ltd.

Dimitrova, M., Vegt, N., & Barakova, E. (2012). Designing a system of interactive robots for training collaborative skills to autistic children. Interactive Collaborative Learning, *Proc. ICL2012*, (pp. 1-8), Villah, Austria, IEEE Xplore®.

Dubal, S., Foucher, A., Jouvent, R., & Nadel, J. (2011). Human brain spots emotion in non humanoid robots.Social, Cognitive and Affective Neuroscience, 6(1), 90-97.

Duquette, A., Michaud, F., & Mercier, H. (2008). Exploring the use of a mobile robot as an imitation agent with children with low-functioning autism. *Autonomous Robots*, 24(2), 147-157.

Ekman, P., Friesen, W.V., & Hager, J.C. (2002). *The Facial Action Coding System. Second edition*. Salt Lake City: Research Nexus eBook, London: Weidenfeld& Nicolson.

Escobedo, L., Nguyen, D., Boyd, L., Hirano, S., Rangel, A., García, D., Tentori, M., & Hayes, G. R. (2012). MOSOCO: A mobile assistive tool to support children with autism practicing social skills in real-life situations. *Proc*. *CHI'12* (pp. 2589-2598), Austin, Texas.

Feil-Seifer, D. & Mataric, M.J.(2012). Distance-based computational models for facilitating robot interaction with children. *Journal of Human-Robot Interaction*, 1(1), 55-77.

Ferrari, P.F., Tramacere, A., Simpson E.A., & Iriki, A. (2013). Mirror neurons through the lens of epigenetics. *Trends in Cognitive Sciences*, 17(7), 450-457.

Galda, H. (2011). Image Processing with Scilaband Image Processing Design Toolbox. Retrieved from: http://forge.scilab.org/index.php/p/IPD/downloads/388/

Haffey, A., Press, C., O'Connell G., & Chakrabarti B. (2013).Autistic traits modulate mimicry of social but not nonsocial rewards. *Autism Research*, August 12, doi: 10.1002/aur.1323, 1-7.

Hafner, V. V. & Kaplan, F. (2005). Learning to interpret pointing gestures: Experiments with four-legged autonomous robots.In S. Wermter et al. (Eds.), *Biomimetic Neural Learning* (pp. 225–234), LNAI 3575.

Happe, F. (1999). Autism: Cognitive deficit or cognitive style? *Trends in Cognitive Sciences*, 3 (6), 216-222.

Hashimoto, S., Ishida, A., Inami, M., & Igarashi, T. (2011). TouchMe: An augmented reality based remote robot manipulation. *Proc*. *21st International Conference on Artificial Reality and Telexistence ICAT11*, (pp. 1-6), Osaka, Japan.

Huskens, B., Verschuur, R., Gillesen, J., Didden, R., & Barakova, E. (2013). Promoting question-asking in school-aged children with autism spectrum disorders: Effectiveness of a robot intervention compared to a human-trainer intervention. *Journal of Developmental Neurorehabilitation*, 16 (5), 345-356.

Johansson, I. (1998). Pattern as an ontological category. In N. Guarino (Ed.) *Formal Ontology in Information Systems FOIS 1998*, (pp. 86-94), Amsterdam, IOS Press.

Johansson, I. (2006). Inference rules, emergent wholes and supervenient properties. *TripleC – Cognition, Communication, Co-operation*, ISSN 1726-670X.

Kandel, E. (2013). The new science of mind and the future of knowledge. *Neuron*, 80, 546-560.

Kim, E. S.,Leyzberg, D., Tsui, K.M., & Scassellati, B. (2009). How people talk when teaching a robot. *Proc. 4th ACM/IEEE International Conference on Human-Robot Interaction HRI* '09 (pp. 23-30), ACM, La Jolla, CA, USA.

LeGoff, D. B. (2004). Use of LEGO(c) as a therapeutic medium for improving social competence. *Journal of Autism and Developmental Disorders*, 34 (5), 557-571.

Liu, Y.,Cherkassky, V.L., Minshew, N.J., & Just, M. A. (2011). Autonomy of lower-level perception from global processing in autism: Evidence from brain activation and functional connectivity. *Neuropsychologia*, 49, 2105–2111.

Milivojevic, B., Clapp, W.C., Johnson, B.W., & Corballis, M. C. (2003). Turn that frown upside down: ERP effects of thatcherization of misorientated faces. *Psychophysiology*, 40(6), 967-978.

Otsu, N. (1979). A threshold selection method from gray-level histogram. *IEEE Transactions on Systems, Man and Cybernetics*, 9, 62-66.

Perlin, K. (1997, August). Layered compositing of facial expression. In *International Conference on Computer Graphics and Interactive Techniques: ACM SIGGRAPH 97* Visual Proceedings: The art and interdisciplinary programs of SIGGRAPH'97 (Vol. 3, No. 08, pp. 226-227).

Rausch, E. (1988). Edwin Rausch: The ontological morphology of Gestalten. In B. Smith (Ed.), *Foundations of Gestalt Theory* (pp. 50-58), Munchen-Vien: PhilosophiaVerlag.

Rizzolatti, G. & Fabbri-Destro, M. (2010). Mirror neurons: from discovery to autism. *Experimental Brain Research*, 200 (3-4), 223-237.

Robins, B., Dautenhahn, K., & Dubowski, J. (2006). Does appearance matter in the interaction of children with autism with a humanoid robot? *Interaction Studies*, 7 (3), 479–512.

Smith, B. (1988). Gestalt theory: An essay in philosophy (pp. 11-81). In B. Smith (Ed.), *Foundations of Gestalt Theory*, Munchen-Vien: PhilosophiaVerlag.

Tentori, M.,& Hayes, G. R. (2010). Designing for interaction immediacy to enhance social skills of children with autism. *Journal of Ubiquitous Computing*, ACM Press, 51-60.

Thompson, P. (1980). Margaret Thatcher: a new illusion. *Perception*, 9, 483-484.

Vechera, S.P., Vogel, E.K., & Woodman G.F. (2002). Lower region: A new cue for figure-ground assignment. *Journal of Experimental Psychology: General*, 131(2), 194-205.

Vekker, L.M.(1980). Thinking and intellect. *Psychological processes*, Volume 2 (In Russian).

Von Ehrenfels, C. (1988). On 'Gestalt' qualities. In: B. Smith (Ed.), *Foundations of Gestalt theory* (pp. 82-117),Munchen-Vien: PhilosophiaVerlag.

Wertheimer, M. (1958). Principles of perceptual organization. In D.C. Beardslee & M. Wertheimer (Eds.), *Readings in Perception* (pp. 121-129), Princeton: Van Norstrand.

White, S., O'Reilly, H., & Frith, U. (2009). Big heads, small details and autism. *Neuropsychologia*, 47, 1274–1281.

Zeigarnik, B. (1955). On finished and unfinished tasks. In Ellis (Ed.), *A Source Book in Gestalt Psychology* (pp. 300-314), Routledge.

Zwickel, J., White. S. J., Senju, A.,&Frith, U. (2011). Exploring the building blocks of social cognition: Spontaneous agency perception and visual perspective taking in Asperger syndrome. *Social Cognitive and Affective Neuroscience*, 6, 564-571.