TALKING HEADS THAT COMMUNICATE BY EYES

Isabella Poggi and Catherine Pelachaud
poggi@uniroma3.it

1. Gaze communication

In this poster we present an analysis of gaze as a communicative behavior, aiming at constructing a 3D agent that exhibits communicative facial and eye actions during a simulated dialog situation.

As a communicative signal, gaze necessarily includes two aspects, a signal and a meaning. The signal is the set of physical features and behavior of eyes in gaze, their muscular actions and physiological state; the meaning is the set of beliefs that gaze communicates.

In the following sections, we first explore the meaning side of gaze, looking for the different kinds of information that can be conveyed by eyes. Then we focus on the signal side, trying to single out the physical components of gaze, and making some hypotheses as to which aspects of the signal correspond to which aspects of the meaning. Then we find the corresponding Action Units (the muscular actions found out by Ekman’s FACS), and finally we propose how to simulate them in creating a synthetic face that communicates by gaze.

2. Talking faces and gaze

Among works on animated faces, only a few have taken eye behavior into account (Takeuchi and Nagao, 1993; Cassell et al., 1994; Thorisson, 1997). This work is based on a research (Pezzato & Poggi, 1998), where a number of face-to-face interactions and mass-media recordings were collected, and single items of gaze were analyzed in terms of both their meanings and their signals.

3. Seeing, looking, thinking, talking

Eyes may have at least four different functions, with only one being a strictly communicative function: we use them for seeing, looking, thinking and communicating.

a. Seeing is to use eyes strictly for vision, that is, for storing information coming in through visual perception. For instance, while for waiting the bus I see three people at the bus stop but I do not pay attention to them. To see has then an intake function, it aims at taking information in, not in outputting anything.

b. Looking is to have the intention of seeing: to direct gaze toward something with the goal of getting visual information about it. For instance, while waiting for the bus I see a handsome man at the bus stop and, while he does not see me, I look at him just to enjoy a beautiful sight. This is also a "receptive function" of eyes in that it is not done in order to produce information in other people.

c. Eyes may also be used while thinking. When I am concentrating I can raise my eyes up, or close my eyes (as in cut-off, see Morris, 1977) or have my eyes "lost in the vacuum". This may simply aim at helping us think better, and if this provides others with some information (the fact that I am thinking) this is, however, only a side effect of the primary goal to help us think; it does not have the goal of letting others know I am thinking. In this case my eyes are not, strictly speaking, communicating.

In other cases, instead, I may display this kind of gaze just for the goal of letting others know I am thinking, and this is a "communicative" use of eyes.

d. Finally, communicating is to use eyes (their actions and movements) for the goal of giving information to somebody. For instance, at the bus a boy steps on my toe and I stare at him to reproach him.

4. Communicating eyes: the world and the mind

Among all the meanings we communicate two broad classes can be distinguished: Information on the World and Information on the Speaker’s Mind (Poggi, 1996). The former class includes all the places, times, objects and events, both concrete and abstract we refer to in our verbal or nonverbal discourse; the latter includes the Speaker's mental states that give rise to or have something to do with ongoing interaction: they are, namely, the Speaker's beliefs, goals, emotions and evaluations.

We can distinguish some subclasses of these meanings and ask which of them may be conveyed through eye communication.
5. Information on the World

Within information on the World, we may mention events and their spatial and temporal location; within events, we may refer to properties of concrete and abstract entities (objects, people, animals, discourses...) and to relations among them. Let us see some examples of gaze bearing these kinds of meanings, and of the specific features of eyes in doing this job.

5.1. Deictic eyes

Eyes can make reference to places and entities located in them: by eyes we can point at things or people in a spatial context, a sort of "deictic gaze". This kind of gaze might be paraphrased as: "I am referring to something in that place", where the something might be both a single entity, like a person or an object, or a whole event.

5.2. Adjectival eyes

Eyes may also have an adjectival function, in that they may mention a small number of physical properties of things: by squeezing eyes we may refer to very small objects, by wide open eyes to very large things. This is a case in which a signal cannot be considered completely arbitrary, while not being iconic either. There is a non-arbitrary, natural link between narrow eyes and small things: here gaze refers to a property of an object by squeezing eyes in the same way one would in adjusting vision to the object dimension.

Examples of this gaze are often found in people telling tales to children, where every expressive device is spontaneously used in order to be better understood and to communicate even the emotional nuances of the tale better. Allusion to dimension may be also used in a metaphorical sense, for instance in talking of something conceptually subtle.


During communication the Speaker, through words, gestures, gaze, posture may communicate information about his/her beliefs, goals, emotions and evaluations, that is, Information on the Speaker’s Mind. The signals devoted to communicating this kind of information may be called Mind Markers (Poggi, 1996). Let us see some Gaze Mind Markers, that is, eyes conveying Information on the Speaker’s Mind (see Table 1).
Table 1  
**WORD AND GAZE MIND MARKERS**

| BELIEFS       | DEGREE OF CERTAINTY | I suppose, of course, no, perhaps | raised eyebrows =  
|               |                     |                                   | I am not sure       
|               | BELIEF RELATIONS    | then, but, because               | inner eyebrows closer =  
|               |                     |                                   | I am sure of what I say |
|               | META-COGNITIVE INFORMATION | I’m trying to remember | eyes up =  
|               |                     |                                   | I am trying to set links |
|               |                     |                                   | eyes down in the corner =  
|               |                     |                                   | I am trying to remember |
| GOALS         | PERFORMATIVE       | I suggest I ask I implore        | fixed stare =  
|               |                     |                                   | I dare you |
|               | SENTENCE PLAN      | Do you remember X? Well...       | eyes away from Addressee =  
|               |                     |                                   | this is the topic |
|               |                     |                                   | eyes towards Addressee =  
|               |                     |                                   | this is the comment |
| DISCOURSE PLAN |                     | In conclusion, Coming to...      | raised eyebrows =  
|               |                     |                                   | I want to speak |
|               |                     |                                   | closer eyebrows =  
|               |                     |                                   | I do not understand |
| CONVERSATION PLAN |                 | Sorry if I interrupt you, hmmm... | raised eyebrows =  
|               |                     |                                   | I want to speak |
|               |                     |                                   | closer eyebrows =  
|               |                     |                                   | I do not understand |
| EMOTIONS      |                     | I’m angry, I am sad              | inner eyebrows raised =  
| EVALUATIONS   |                     | handsome tasty bossy              | I am sad |
6.1. Beliefs
The Speaker may provide three types of information about one’s beliefs: degree of certainty, relationships between beliefs and metacognitive information.

6.1.1. Epistemic eyes
While communicating we often mark how reliable the information we are providing is, how certain we are of it. We mark if it is sure, only likely or very unlikely, in verbal language through lexical devices (words like maybe, of course) or grammatical devices (verbal modes like subjunctive or conditional). But to communicate we are not sure of something we also raise our eyebrows.

6.1.2. No abstract relationship through eyes
In words, the relationship between beliefs may be conveyed by causal or adversative conjunctions (because, but) and by gestures stating a link (say, Italians rotate the wrist with curved thumb and index). But this kind of information cannot be conveyed, we think, by eyes.

6.1.3. Metacognitive eyes: the gaze of thought
 Sometimes while speaking we also meta-inform about the source or cognitive status of information we are talking of. The interjections Oh! means "This belief is new to me, I did not know it before"; Well means: "I am still deciding on what to do or to say". Eyes can also be used to provide information about our processes of thought. By eyes we may inform that we are thinking, or trying to remember. This kind of gaze is communicative, sometimes even doubly communicative, at both the literal and the indirect levels of meaning. At the first level, eyes simply inform I am thinking, but at a second level (Poggi, 1997) they may imply: "Let me think, please, let me finish expressing my thoughts", thus being a turn-taking request. Another way to communicate we are thinking is to look down obliquely, which seems to mean not simply "I am thinking now", but "I am trying to remember, to retrieve from memory"; and it may also indirectly ask: "Please let me think by myself", "let me think alone" or, even "I would like to escape from this situation".

6.2. Goals
Within information on the Speaker's goals, we may distinguish information about the goal of a single communicative act (the performative of a sentence) from information on a whole hierarchy of goals, namely the planning of a sentence (sentence goals), or of a monologic discourse (metadiscursive goals), or of the overall arrangement of conversation (metaconversational goals), that is turn-taking and back-channelling.

6.2.1. Performative eyes
Gaze has a performative function when it communicates the illocutionary force of a speech act. A peremptory order is communicated by a strict, serious gaze, with inner parts of eyebrows slightly closed as in an angry face (Poggi and Pelachaud, 1998). But we also may have an imploring gaze, a defiant gaze, a reproaching gaze and so on.

6.2.2. Topic-comment eyes
In any sentence I utter, a topic and a comment can be distinguished: the topic is information I take for granted and shared by me and my interlocutor, the comment being information I consider my new and relevant contribution to ongoing discourse. The comment part of a sentence can typically be stressed by opening eyes wide and raising eyebrows, while clearly and ostentatiously directing one’s gaze to one’s interlocutor.

6.2.3. Meta-conversational eyes: turn-taking and back-channelling
Another well-known function of gaze behavior is in regulating conversation. Two kinds of goals are prominent here: a. turn allocation ("I want to speak now", or "Please speak now"); b. back-channelling ("I am following / not following you", "I agree/disagree with you"). Simply gazing at a conversant is a way to pass speaking turn; while asking for a speaking turn is better done by wide opening eyes, like in breathing to start speaking.
Among back-channelling goals, frowning signals that I do not understand what you're saying, or, on the indirect meaning level, that I do not agree with you. With signals of positive understanding and agreement, instead, head nods seem more relevant than eye movement: a head nod is more evident than the accompanying eye-blink.

6.2.4. No meta-discoursive eyes

In verbal, but also in gestural communication, many Mind Markers are devoted to making clear at what point of the discourse plan the Speaker presently is. They are verbal expressions like: in conclusion, to sum up, or I am now coming to the second point, and so on. Yet, eyes perhaps cannot bear information about relationships between beliefs. We think it is not in principle possible for eyes to indicate relationships and, moreover, such abstract metatextual relationships.

6.3. Emotion and evaluation through eyes

One kind of information that eyes typically bear are emotions. Our gaze may show social emotions, those one can only feel towards another person, like love, admiration, scorn, anger at somebody; or else individual emotions, eventually triggered by natural events but not directed toward anyone in particular: fear, terror, joy, sadness, surprise, excitement, worrying, dismay. We do not think, instead, that specific eye actions communicate evaluation per se; eyes do have an evaluative import, but only when intertwined with other kinds of information, for instance performatives and emotions. For instance, reproaching and scornful eyes communicate a negative evaluation.

7. The signal of gaze

Is there an alphabet of eyes? Can we decompose each single gaze into a combination of parts or aspects comparable to phonemes or distinctive features in a verbal language? Our hypothesis is that any single gaze can be analyzed in terms of a small set of physical parameters like eye direction, humidity, eyebrow movements and the like, each of which may be attributed a small set of values: the combination of those values (one value for each parameter) provides a precise description of any gaze under analysis. In other words, we have trying to do what Stokoe (1978) did for Sign Language: to find out the formational parameters of gaze.

7.1. The formational parameters of gaze

Using videotaped data, we have singled out some parameters (muscular actions and physiological states of the eye region) in terms of which we analyzed gaze (Table 2). Some of these parameters can be realistically used to simulate a talking face because they correspond to specific Action Units (AUs) of Ekman's FACS (Table 3). Each AU describes one or more actions of one or more specific muscles; by simulating these AUs through computer graphics techniques it is possible to build an animated face that exhibits communicative gaze behavior. The anatomical portion of the face we take into account includes the following parts: eyebrows, upper eyelids, eyes, lower eyelids, and wrinkles. Within each part, different subparts, aspects or actions are considered relevant, or just their presence / absence: for instance, eyebrows are subdivided into inner, medial and outer part (after Ekman, 1979; Ekman and Friesen 1978), even if sometimes they do not correspond to AUs by themselves but only in reciprocal combination or with other parts.
Table 2.

1. EYEBROWS:
right/left: inner part: up / down/ central
   medial part: up / down/ central
   outer part: up / down/ central

2. EYELIDS:
right / left:
   upper: default / raised / lowered
      default / tense / corrugated
      blinking / winking / closed

   lower: default / raised / lowered
      default / tense / corrugated

3. WRINKLES:
   vertical / horizontal / curved / oblique
   central / lateral / all along forehead / between brows
   crow's feet / bulging (lower lid) / bagging (lower lid)

4. EYES:
right / left
   humidity: dry / wet / tears
   reddening: default / reddened
   pupil dilation: default / dilated / narrow
   direction of the head: forward / up / down / left / right / backward
   eye movements: forward / left / right/ up / down
Table 3

1. EYEBROWS:
right/left eyebrows
AU1: inner and central part of eyebrow UP
AU2: outer part of eyebrow UP
AU4: inner part / inner and central parts / inner, central and outer parts of eyebrow DOWN
AU1 + AU2: inner, central and outer parts of eyebrow UP
AU1 + AU4: inner and central parts of eyebrow UP and CENTRAL
AU1 + AU2 + AU4: inner, central and outer parts of eyebrow UP and CENTRAL

2. EYELIDS:
left / right eyelids
AU5: upper eyelid RAISED
AU41: upper eyelid LOWERED
AU42: upper eyelid (very) LOWERED
AU43: upper eyelid CLOSED
AU6: upper eyelid (LOWERED and CORRUGATED) + lower eyelid (RAISED and CORRUGATED)
AU7: upper eyelid (LOWERED and TENSE) + lower eyelid (RAISED and TENSE)
AU44: upper eyelid LOWERED + lower eyelid ((very) RAISED and (very) TENSE)
AU45: BLINK
AU46: WINK

3. WRINKLES:
AU1: VERTICAL / OBLIQUES between brows
AU2: HORIZONTAL lateral part of forehead
AU4, AU1 + AU4, AU4 + AU5: VERTICAL between brows
AU1 + AU4, AU1 + AU2 + AU4: CURVED central part of forehead + OBLIQUE between brows
AU1 + AU2: HORIZONTAL all along forehead
AU7, AU5 + AU7, AU44: BULGE lower lid
AU6, AU6 + AU43: CROW’S FEET

4. EYES
direction of the head
AU51: LEFT
AU52: RIGHT
AU53: UP
AU54: DOWN
AU55: TILT LEFT
AU56: TILT RIGHT
AU57: FORWARD
AU58: BACKWARD
eye movements
AU61: LEFT
AU62: RIGHT
AU63: UP
AU64: DOWN
AU65: walleye
AU66: crosseye

7.2. The relevance of gaze parameters for communication
Why are these parameters relevant? **Eyebrows** are typically engaged in the expression of emotions like fear, anger, surprise, worrying (Ekman, 1979), but also in greetings (Eibl-Eibesfeldt, 1974) and in topic-comment marking and emphasis (Cassel & Prevost, 1996; Torres.et.al., 1997). **Eyelids** are important because they determine the openness of eyes, thus marking the withdrawing from interaction in cut-off, underlining excitement in flirting and so forth. As for **eyes**, humidity may be relevant in joy or enthusiasm (bright eyes), or in sorrow (tears); reddening may be a cue to crying (and then sadness) or to rage (bloodshot eyes). Pupil dilation is a cue to excitement or other kinds of arousal. In the eyes’ spatial behavior, we must take into account the reciprocal relationships among eyes, head and trunk, and their relationship to where the interlocutor is.

8. From meanings to gaze

As was shown above, gaze may provide many types of information, both on the World and on the Speaker’s beliefs, goals and emotions. Now, all of these **semantic contents** may be represented in terms of **“COGNITIVE UNITS”**, declarative representations of semantic primitives, by which all kinds of semantic content, including communicative intentions, word meanings and emotions, may be represented.

For instance, the performative of **imploration**, which is the specific goal of a communicative act, is represented in the Speaker’s mind as containing the following cognitive units:

- the Speaker S asks the Listener L to do something
- which is in the interest of S,
- and while acknowledging L’s power over S.

On the meaning side, then, an imploring gaze may be represented in terms of these cognitive units; while on the signal side it is performed by AUs which move the Speaker’s eyebrow inner parts up (Poggi & Pelachaud 1998). The **comment** of a sentence may be represented like this:

- what I am saying at the moment
- is the most important part of my sentence

This is communicated by gazing at the interlocutor straight in his/her eyes, and/or raising the eyebrows.

Finally, an emotion of **anger** may be represented as:

- somebody has thwarted a quite important goal
- I have the goal of aggressing him/her

and it is expressed by inner parts of eyebrows closer and vertical wrinkles between them.

At the extent to which it is possible to state a correspondence between **specific clusters of cognitive units** and **specific clusters of AUs**, drawn from the presented parameters, we may endow a talking face with a library of communicative gaze.
9. System overview

Here is an overview of our system.

1. The **input** of the system is a **sequence of sentences** to be uttered in subsequent turns, whose meanings (propositional meaning, performatives, topic-comment structure, emotions felt by the Speaker...) are **represented in terms of cognitive units**.

2. The **system computes** the appropriate **gaze parameter values** chosen from the library of gaze communicative behavior, **and the corresponding AUs**.

3. These values are given to a **graphics engine** that generates the animation of 3D talking heads.

4. The system then generates gaze patterns that enhance the comment of the sentence, ask for speaking turn, express surprise or anger, show comprehension or disagreement, and so on.