TR－H－022

Indirect Measurement of Feature Saliency in Face Processing

Takashi KATO　Masaomi ODA

1993. 8. 17

ATR人間情報通信研究所

〒619-02 京都府相楽郡精華町光台2-2　☎07749-5-1011

ATR Human Information Processing Research Laboratories

2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02 Japan

Telephone: +81-7749-5-1011
Facsimile: +81-7749-5-1008

Indirect Measurement of Feature Saliency in Face Processing[1]

Takashi Kato and Masaomi Oda

ATR Human Information Processing Research Laboratories
2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02, Japan
Email: tkato or oda@hip.atr.co.jp

ABSTRACT

This article introduces a newly devised orienting task and discusses its utility in face processing research. As an example, we present the experimental results obtained in a pilot study investigating feature saliency. In this study, subjects were presented with a series of line-drawn faces and were asked to select 10 of their favorite faces or 10 faces that were most similar to a particular target. The line-drawn faces were randomly selected from a pool of nearly 60,000 different faces, which were drawn by choosing one of 3 possible values for each of 10 facial features (i.e., 3^{10} possible faces). Assuming that the subjects would select the faces with matching values for important features, the relative importance of facial features may be assessed by comparing the amount of variance "tolerated" by the subjects for each feature. It was found, for example, that in similar face retrieval, the variance was significantly smaller for eyebrow tilt, eye shape, and face shape than for other features, such as eyebrow position, eye position, or nose length. The utility of the present orienting task was discussed in terms of what such experimental results reveal about feature saliency.

[1] Thanks are due to Yoh'ichi Tohkura and Shigeru Akamatsu for their support and encouragement throughout the project reported here.

INTRODUCTION

A taxonomy of experimental methods for feature saliency

Feature saliency in face processing has been investigated using a variety of experimental methods, which may be classified into five categories according to the types of responses required of subjects: Recall, construction, recognition, semantic judgment, and perceptual judgment.

Ellis, Deregowski, and Shepherd (1975), for example, utilized a recall method in which subjects were asked to freely describe target faces. The frequency of mention of different features was noted and used as a basis of assessing the relative importance that the subjects attached to different features of the faces. The assumption here is that important features are mentioned more frequently than less important features. A more direct method of recall was used by Laughery, Alexander, and Lane (1971) who asked their subjects, at the conclusion of the recognition experiments, to state which facial features they regarded as important in making identification decisions.

The construction method requires subjects to reproduce a target face with the assistance of a composite-face generation system, such as the Photofit system or the Identikit system. For example, Ellis, Shepherd, and Davies (1975) used the Photofit system, which consists of a number of interchangeable variants of five features, and asked their subjects to reconstruct a target face that was itself constructed using the Photofit system. The relative importance of these five features defined by the system was assessed by noting the likelihood that the correct variant of the feature was successfully selected by the subjects.

In the recognition method, subjects are asked to recognize an either semantically or episodically familiar face with one or more features masked, or to identify a target face that is mingled with distractors some of whose features have been replaced with a variant. Haig (1985), for example, used what he called the distributed aperture technique, which enabled the construction of a set of stimulus faces that differed according to which facial feature was masked and to what extent. In each trial, subjects were presented with a target face, followed by an array of four faces from which they were to choose one that they judged had previously been presented as a target. The percentage of correct responses as a function of which face portions were masked was used to assess the saliency of facial features.

Roberts and Bruce (1988) utilized both the recognition and semantic-judgment methods to obtain information about the relative importance of internal features in the familiarity and sex judgments of faces. Subjects were presented with a series of faces in which either the eyes, nose, mouth, or none of the above was masked and were asked to decide as quickly as possible whether or not the presented face was familiar to them, or whether the presented face was that of a male or female. The reaction times as a function of the features masked were the basis of assessing the relative importance of the internal features in the familiarity and sex judgments.

In the perceptual-judgment method, subjects are asked to make decisions on perceptual properties of test faces, some of which are made different from target faces by omitting or altering particular features. The saliency of facial features is analyzed based on the speed and accuracy of the subjects' responses. Matthews (1978), for example, asked his subjects to make same-different judgments on simultaneously presented pairs of faces, which were constructed using the Identikit system. The number and types of features that differed between each pair were manipulated and reaction times and detection errors were analyzed with respect to these experimental manipulations. In the Fraser, Craig, and Parker (1990) study, subjects were successively presented with the parts of a schematic face, and asked to indicate whether successively presented stimuli showed a complete or incomplete face. This is an example of the perceptual-judgment method using the omission technique where one of the face parts could be omitted in some trials. The study also used the recognition method and the substitution technique such that subjects were asked to decide whether or not successively presented stimuli comprised a previously presented target face when in some trials, one of the original face parts was substituted with its counterpart from a different face. Reaction times and detection errors were analyzed as a function of which feature was omitted or substituted.

Face retrieval as an orienting task

In this article, we introduce yet another experimental method for investigating feature saliency in face processing. An important characteristic of this new "database retrieval method" is its indirect or implicit approach to the measurement of feature saliency. Subjects are induced to engage in a database retrieval task, which requires them to retrieve (or collect) a number of facial images that satisfy a particular retrieval goal. It may be said that in this method, the effect of feature manipulation is indirectly achieved by requiring subjects to view and evaluate a number of different faces before deciding the final set of faces that meets their selection criteria. In other methods that use the feature omission, substitution, or masking technique, the experimental manipulation of feature characteristics is more explicit, in that subjects are explicitly instructed to attend to the possibility that such feature alteration is present in the presented stimuli (e.g., forced-choice recognition, perceptual judgment).

Another important characteristic of the database retrieval method is that the total number of faces in the database is made so large that it wold be impractical for subjects to view all the faces in the database. This implies that the subjects would need to make some sort of compromise in their effort to retrieve the best faces, which are necessarily different from each other. A primary interest would be to find any difference in the degree to which the subjects might compromise on different features of the faces. For example, it should be possible to assess the saliency of facial features by comparing the observed variances of the feature values among those faces collected by the subjects. Further, if one feature shows a smaller variance than do other features, this fact may be taken as evidence that this feature is considered more important than the others.

The purpose of this article is to demonstrate the utility of the database retrieval method for face processing research. To this end, we will illustrate the results of an experimental evaluation of a facial image retrieval system.

METHOD

Subjects

A total of 40 Doshisha University students, 13 male and 27 female, was recruited and paid to participate in the experiment.

Material and apparatus

The facial-image retrieval system used in the present study was developed for the initial investigation of our previously proposed "context-driven retrieval mechanism" (e.g., Oda, 1991). Line-drawn faces were used as image data to simplify database construction and manipulation of experimental variables.

The system contains nearly 60,000 line-drawn faces, each of which is drawn by choosing particular values for 10 facial features. The facial features used are: Eyebrow position (BP), eyebrow tilt (BT), eye position (EP), eye shape (ES), face shape (FS), mouth position (MP), mouth length (ML), nose position (NP), nose length (NL), and ear position (RP). There are three possible values for each feature, thus producing 59,049 (i.e., 3^{10}) possible faces. Examples of line-drawn faces are shown in Figure 1.

This facial-image retrieval system simultaneously presents 10 line-drawn faces on a computer display, from which the user selects the face or faces that meet his or her selection criteria. These user-selected faces are then stored in a background buffer. With the context-driven retrieval mechanism, the system attempts to present to the user those faces that are calculated to be most similar to the user-selected faces in the buffer. With the context-driven retrieval mechanism turned off, the system presents 10 faces that are randomly chosen from the total database. The system can hold up to 10 of the most recently selected faces in the background buffer, which the user can at anytime call up and re-evaluate the currently-held faces and decide whether to retain or discard any of them.

Design and procedure

There were two different retrieval tasks. In the favorite-face retrieval task, one group of subjects (6 male and 14 female) was asked to retrieve 10 of their favorite faces from the database. In the similar-face retrieval task, another group of subjects (7 male and 13 female) was asked to retrieve 10 faces that they judged were most similar to a particular target face.

In each retrieval, the subjects completed the task once with the assistance of the context-driven retrieval mechanism and once without it. Half of the subjects first attempted context-driven retrieval and the other half first attempted random retrieval. The order of these two modes of retrieval was randomized across the subjects. Thus, the type of retrieval tasks (i.e., favorite and similar face retrieval) was manipulated between subjects, and the modes of retrieval (i.e., with or without the context-driven retrieval mechanism) and facial features within subjects.

The subjects were randomly assigned to one of four conditions (2 retrieval tasks x 2 orders of retrieval modes). After receiving retrieval task instructions and an explanation of the facial image retrieval system, the subjects were given an opportunity to familiarize themselves with the system and mouse operation. The practice session lasted five minutes. Following the practice period, the subjects attempting similar face retrieval were given five more minutes to select their favorite face. A printed copy of that face was then given to them as their retrieval task target.

In all retrieval cases, the subjects were asked to continue the retrieval task until they were satisfied with the 10 faces that they collected in the buffer. At the end of each retrieval session, the subjects were asked to identify the most satisfactory face from among the 10 selected faces and to orally explain why they judged it to be the best among the favorite or similar faces that they had collected.

RESULTS

Since the main purpose of this article is to discuss the database retrieval method as a viable alternative for investigating feature saliency in face processing, our focus was on analyzing the data obtained in the random retrieval mode. A comparison of context-driven and random retrieval was presented in Kato and Oda (1993) in which the effectiveness of the context-driven retrieval mechanism was extensively examined.

The subjects took on average 7 minutes and 32 seconds to complete the retrieval of favorite faces, and 9 minutes and 13 seconds to complete the retrieval of similar faces. The mean total number of faces examined by the subjects was 479 for favorite face retrieval, and 792 for similar face retrieval. The results suggest that the subjects in similar face retrieval took less time to examine presented faces. However, caution needs to be exercised in interpreting these results, since the task completion time included the time the subjects spent examining the selected faces in the buffer, and the amount of the buffer viewing was at their discretion.

Variance data

As mentioned before, each subject produced a set of 10 faces as the outcome of a given retrieval task. We calculated, for each of the 10 facial features, the variance of the feature parameter values among the 10 faces in the final set produced by each subject. The variance data as a function of retrieval tasks and 10 facial features are shown in Figure 2.

A 2 (orders of retrieval modes) x 10 (facial features) analysis of variance (ANOVA) conducted on the variance data in favorite face retrieval indicated that there were significant differences in the mean variances among the 10 facial features, $F(9, 162)=9.50$, $p<0.0001$. The presence or absence of the prior retrieval experience with the context-driven retrieval mode was not significant, $F(1, 18)=1.83$, $p>0.19$, nor was the interaction between facial features and prior experience, $F<1$.

A Tukey pairwise comparison ($\alpha=0.05$) showed that face shape and eyebrow tilt had significantly smaller variance than did nose length, nose position, eyebrow position, mouth length, and ear position. The variance of face shape was significantly smaller than that of mouth position, too. We also found that the variance of eye shape was significantly smaller than that of nose length, nose position, and eyebrow position, while the variance of eye position was significantly smaller than that of nose length and nose position.

A 2 (orders of retrieval modes) x 10 (facial features) ANOVA applied to the variance data in similar face retrieval showed that while the main effect of facial features was significant, $F(9, 162)=22.42$, $p<0.0001$, the main effect of the presence or absence of the prior retrieval experience with context-driven retrieval did not reach the significance level, $F(1, 18)=3.87$, $p>0.06$. There was no significant interaction between facial features and prior retrieval experience, $F<1$.

A Tukey pairwise comparison ($\alpha=0.05$) indicated that eyebrow tilt, face shape, and eye shape had significantly smaller variance than the remaining seven features, and that mouth position, mouth length, and eye position had significantly less variance than nose position.

A 2 (retrieval tasks) x 10 (facial features) ANOVA applied to the combined data of favorite and similar face retrieval indicated that the main effects of retrieval tasks and facial features were both significant, $F(1, 38)=15.20$, $p<0.0005$, for retrieval tasks, and $F(9, 342)=29.43$, $p<0.0001$, for facial features. The interaction between retrieval tasks and facial features was also significant, $F(9, 342)=2.72$, $p<0.005$.

The simple main effects of retrieval tasks were tested for each of the 10 facial features, using multiple t tests. Each t test was conducted at the 0.01 level of significance. It was found that the difference in variance between favorite and similar face retrieval was significant for eyebrow tilt and eye shape. The difference was found to be marginally significant ($0.01<p<0.05$) for face shape.

Hit-rate data

The goal of similar face retrieval was to collect a set of 10 faces that were most similar to a particular target face. We calculated, for each of the 10 facial features, the hit rate between the target face and the 10 faces in the final set produced by each subject. The hit rates as a function of retrieval tasks and 10 facial features are shown in Figure 3.

A 2 (orders of retrieval modes) x 10 (facial features) ANOVA showed that hit rates were significantly different among the features, $F(9, 162)=21.70$, $p<0.0001$. However, neither the presence/absence of the prior retrieval experience with context-driven retrieval ($F(1, 18)=3.14$, $p>0.09$), nor the interaction between facial features and prior retrieval experience ($F<1$) was found to be significant.

A Tukey pairwise comparison ($\alpha=0.05$) revealed that eyebrow tilt had a significantly higher hit rate than did the other features, except for face shape. Face shape had a significantly higher hit rate than the eight remaining features, except for eye shape. The hit rate of eye shape was significantly higher than that of the seven remaining features, except for mouth position. The hit rate of mouth position was higher than that of nose length and ear position.

Unlike similar face retrieval, there was no externally given target face in favorite face retrieval. However, at the conclusion of the retrieval task, the subjects were asked to specify the face that they judged to be the best among the 10 favorite faces stored in the final set. We calculated, for each of the 10 facial features, the frequency of occurrence of the feature value commonly shared between the best face and the remaining faces. In the remainder of this article, we refer to this frequency as a hit in favorite face retrieval.

A 2 (orders of retrieval modes) x 10 (facial features) ANOVA showed that the main effect of facial features was significant, $F(9, 162)=7.06$, $p<0.0001$. However, neither the prior retrieval experience ($F<1$) nor the interaction effect ($F<1$) was significant.

A Tukey pairwise comparison ($\alpha=0.05$) indicated that eyebrow tilt had a significantly higher hit rate than did the other features, except for face and eye shape. The hit rate of face shape was significantly higher than the other features, except for eyebrow tilt, eye shape and eye position.

Although the hits in similar and favorite face retrieval are not exactly comparable, a 2 (retrieval tasks) x 10 (facial features) ANOVA was applied to the combined hit-rate data of similar and favorite face retrieval. The results showed that the main effects of retrieval tasks and facial features were both significant, $F(1, 38)=9.30$, $p<0.005$, for retrieval tasks, and $F(9, 342)=26.69$, $p<0.0001$, for facial features. The interaction between retrieval tasks and facial features was also found to be significant, $F(9, 342)=3.34$, $p<0.001$.

The simple main effects of retrieval tasks, conducted at the 0.01 level of significance using multiple t tests, revealed that the difference in the hit rates between similar and favorite face retrieval was significant only for eyebrow tilt. A marginally significant difference ($0.01<p<0.05$) was found for face and eye shape.

Explicit and Implicit Features

The facial features so far discussed refer to the characteristics of individual features, which we might call "explicit features." There are also features of a rather implicit nature, such as spatial relations between individual features, which we might call "implicit features." We attempted a preliminary investigation of implicit features by defining six such features: the positional distances between eyebrow and eye (B-E), eyebrow and nose (B-N), eyebrow and mouth (B-M), eye and nose (E-N), eye and mouth (E-M), and nose and mouth (N-M). We classified the parameter values of these implicit features into the same three categories as the explicit features so as to make their value scales comparable. The mean variances of explicit and implicit features in favorite and similar face retrieval are shown in Figure 4.

A 2 (explicit and implicit features) x 2 (favorite and similar face retrieval) ANOVA indicated that while the main effect of the types of features was significant, $F(1, 38)=45.83$, $p<0.0001$, the main effect of retrieval tasks was non-significant, $F<1$. The interaction between types of features and types of retrieval tasks was significant, $F(1, 38)=16.88$, $p<0.0005$.

A 2 (retrieval tasks) x 6 (implicit features) ANOVA indicated that the main effect of implicit features was significant, $F(5, 190)=2.73$, $p<0.05$, and that the interaction between retrieval tasks and implicit features was marginally significant, $F(5, 190)=2.03$, $0.05<p<0.1$. The main effect of retrieval tasks was not significant, $F(1, 38)=1.69$, $p>0.1$. Figure 5 shows the variance data as a function of retrieval tasks and five implicit features.

Verbal reports

As mentioned in the Method section, at the conclusion of each retrieval task, the subjects were asked to select the best face from among the 10 faces in the final set and give the reason for their choice. These reasons were classified into three categories depending on whether they referred to physical characteristics of individual features (e.g., big eyes), relations between features (e.g., space between eyebrows and eyes), or emotional aspects of the face (e.g., gentle). The frequency of occurrence in each category was then counted. If more than one category was given to the same face by the same subject, the frequency was divided by the number of categories. The percentages of mention as a function of retrieval tasks and selection reasons are shown in Figure 6. The results indicated that the subjects tended to refer more to emotional aspects of the face after favorite face retrieval and more to individual features of the face after similar face retrieval.

DISCUSSION

In the present experiment, the subjects were presented with line-drawn faces, 10 at a time, and were asked to produce a set of 10 faces that satisfied a particular retrieval goal (i.e., retrieving favorite faces or faces similar to a given target face). These faces were supposed to be selected from the database of nearly 60,000 different faces under the condition that all presented faces were randomly chosen by the system. We expected that under such circumstances, the subjects would have to make some compromise in their selection criteria so that they could conclude a given retrieval task within a reasonable period of time.

Two immediate questions were whether such compromise might be made evenly or unevenly across different facial features, and how the degree of such compromise for each feature might be measured so as to determine the relative importance, if any, of different features. One useful measure of the degree of compromise in this regard seemed to be the variance, calculated for each facial feature, of the feature's parameter values among the final set of 10 faces produced by each subject. The rationale was that if any facial features were more important than others, the subjects would select the faces with matching parameter values for these important features, thus resulting in a smaller variance for the more important features than the less important ones.

Saliency of facial features

It is clear from the analyses of the variance data that in favorite face retrieval, the most salient features of the present line-drawn faces are face shape and eyebrow tilt, closely followed by eye shape and eye position. In similar face retrieval, whereas eyebrow tilt, face shape, and eye shape are clear winners, mouth position, mouth length, and eye position appear to be a distant second. The results in similar face retrieval appear to be consistent with the previous studies showing the saliency of face shape or head outline (e.g., Haig, 1985; Fraser *et al*., 1990) and the eyebrow-to-eye area (e.g., Haig, 1986), and the secondary saliency of the mouth area (e.g., Davies, Ellis, & Shepherd, 1977).

Although the overall variance was significantly higher in favorite face retrieval than in similar face retrieval, this effect was mostly due to the much reduced variance for eyebrow tilt and eye shape in similar face retrieval. One possible interpretation of these results is that the variants of eyebrow tilt and eye shape might be more permissible in favorite face retrieval where different shapes of these features could be seen as forming different "expressions" of the same, favorite face. Another possibility is that the subjects in favorite face retrieval might have collected more than one type of favorite face, thus increasing the overall variance in favorite face retrieval. This interpretation, however, does not explain why the significant increase in the total variance in favorite face retrieval resulted mostly from the increased variance in eyebrow tilt and eye shape. One might say that eyebrow tilt and eye shape might have played a minor role in defining favorite faces. This interpretation does not seem tenable, since in favorite face retrieval, these two features nevertheless showed a significantly smaller variance than the other features.

In similar face retrieval, the subjects were engaged in a task that required them to find the faces that were most similar to a given target face. Therefore, the saliency of a given feature needed to be evaluated not only by the amount of variance but also by the degree of correspondence (i.e., hits) in its parameter value between the user-selected faces and the target face. As expected, those features identified as salient because of their smaller amount of variance did show higher hit rates than did the other non-salient features.

Taken together, it seems clear that the subjects in the present experiment made a clear choice of salient features, given the task demands such that they were to collect 10 faces of a defined nature from those faces randomly chosen from the pool of nearly 60,000 different faces. It is worth noting that eyebrow tilt, face shape and eye shape seem to have been treated more importantly in similar face retrieval than in favorite face retrieval. It is also interesting to note that the subjects in similar face retrieval tended to refer to physical characteristics of faces (e.g., big eyes) in reporting the reason for choosing a particular face as the most similar face, while the subjects in favorite face retrieval tended to describe their most favorite face by referring to emotional aspects of the face (e.g., looks sharp). It seems important that the issue of feature saliency in face processing is viewed in the context of "task-appropriate processing."

The database retrieval method

As shown above, the database retrieval method introduces a matrix of subject-by-feature variance data (i.e., the variance of each individual feature from each subject) as a basis of determining the relative importance of facial features. This is a new addition to the repertoire of useful measures in face processing research, which includes eye-movement recording (e.g., Walker-Smith, Gale, & Findlay, 1977) and similarity rating (e.g., Hosie, Ellis, & Haig, 1988) as well as more popular measures of speed and accuracy such as reaction time and response error. The variance data introduced by the database retrieval method is unique in that it is an index of fluctuations that an individual subject allows in completing a given face processing task.

The database retrieval method and the construction method described in the Introduction may be classified as "production"-type methods, which require subjects to produce a face or faces that satisfy a particular retrieval requirement. The subjects in the construction method are required to construct a composite face by choosing appropriate variants of the face parts defined in the construction kit, whereas the subjects in the database retrieval method are required to select appropriate (whole) faces from among the faces presented to them by the database system. The data of main interest in these production-type methods are those characteristics associated with the resulting faces, which are the products of a series of perceptual/cognitive judgments made by the subjects. The perceptual/cognitive judgments themselves, however, are normally unavailable.

The production-type methods may be contrasted with "decision"-type methods, such as the recognition, semantic judgment, and perceptual judgment methods, which require subjects to make overt judgments on the presented faces or face parts as to whether they are identical or similar to defined targets. In such decision-type methods, the perceptual/cognitive judgments themselves are subjected to analysis. In this sense, it can be said that the measurement of feature saliency is more direct in decision-type methods than in production-type methods.

The indirectness of measuring feature saliency may be seen as a weakness of the database retrieval method. However, the unique strength of the database retrieval method comes from the very fact that the analysis is focused on the outcome of the perceptual/cognitive judgments required of subjects in a particular database retrieval task. The research strategy here is to let subjects work out demanding tasks and then compare the parameter values of the processing outcomes in the light of different task demands to see if and where they make any adjustment. A successful application of the database retrieval method, however, depends on the fundamental requirement that sufficient information on the feature parameter values of each datum is available.

For this reason, we developed a set of tools that allow us to define implicit features on a computer display and then automatically calculate the parameter values of these implicit features for the user-selected faces obtained in the experiment. Now we can obtain full information on the parameter values not only of individual facial features (i.e., explicit features) but also of implicit features such as the distance between the tip of the nose and the mouth, the area defined by the tip of the nose and the ends of the mouth, and so forth. Preliminary analysis of the implicit features presented in the Results section indicates a possibility that explicit and implicit features (or first-order and second-order features in the terms of Rhodes, 1988, though face shape is classified as a second-order feature by Rhodes) might interact with retrieval tasks in an intricate manner. We expect to expand the scope and depth of analyses on a variety of implicit features with the aid of these tools.

We are also developing a similar set of tools, which will allow us to specify a variety of explicit and implicit features of real-life, photographed faces on a computer display and then automatically calculate and store the metrics of these features. We expect that these sets of tools will enhance the strength of the database retrieval method and shed further light on the problem of feature saliency in face processing.

REFERENCES

Davies, G, Ellis, H., & Shepherd, J. (1977). Cue saliency in faces as assessed by the 'Photofit' technique. *Perception*, *6*, 263-269.

Ellis, H. D., Deregowski, J. B., & Shepherd, J. W. (1975). Descriptions of white and black faces by white and black subjects. *International Journal of Psychology*, *10*, 119-123.

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1975). An investigation of the use of the Photofit technique for recalling faces. *British Journal of Psychology*, *66*, 29-37.

Fraser, I. H., Craig, G. L., Parker, D. M. (1990). Reaction time measures of feature saliency in schematic faces. *Perception*, *19*, 661-673.

Haig, N. D. (1985). How faces differ — a new comparative technique. *Perception*, *14*, 601-615.

Haig, N. D. (1986). Investigating face recognition with an image processing computer. In H. D. Ellis, M. A. Jeeves, F. Newcombe, & A. Young (Eds.), *Aspects of face processing* (pp. 410-425). Dordrecht: Nijhoff.

Hosie, J. A., Ellis, H., & Haig, N. D. (1988). The effect of feature displacement on the perception of well-known faces. *Perception*, *17*, 461-474.

Kato, T., & Oda, M. (1993). *Context-driven retrieval and saliency of facial features* (Tech. Rep. No. TR-H-019). Seika-cho, Kyoto: ATR Human Information Processing Research Labs.

Laughery, K. R., Alexander, J. F., & Lane, A. B. (1971). Recognition of human faces: Effects of target exposure time, target position and type of photograph. *Journal of Applied Psychology*, *51*, 477-483.

Matthews, M. L. (1978). Discrimination of Identikit constructions of faces: Evidence for a dual processing strategy. *Perception & Psychophysics*, *23*, 153-161.

Oda, M. (1991). Context dependency effect in the formation of image concepts and its application. Proceedings of the 1991 IEEE International Conference on Systems, Man, and Cybernetics (pp. 1673-1678). Charlottesville, VA.

Rhodes, G. (1988). Looking at faces: First-order and second-order features as determinants of facial appearance. *Perception*, *17*, 43-63.

Roberts, T., & Bruce, V. (1988). Feature saliency in judging the sex and familiarity of faces. *Perception*, *17*, 475-481.

Walker-Smith, G. J., Gale, A. G., & Findlay, J. M. (1977). Eye movement strategies involved in face perception. *Perception*, *6*, 313-326.

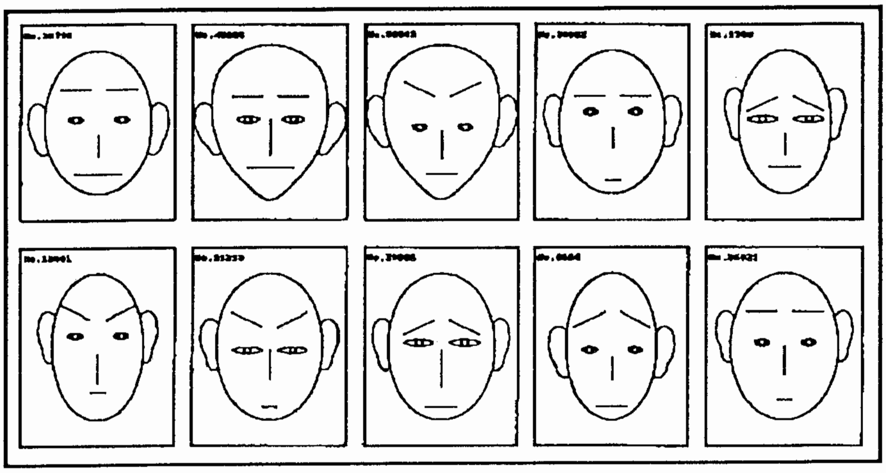

Figure 1. Examples of Line-Drawn Faces

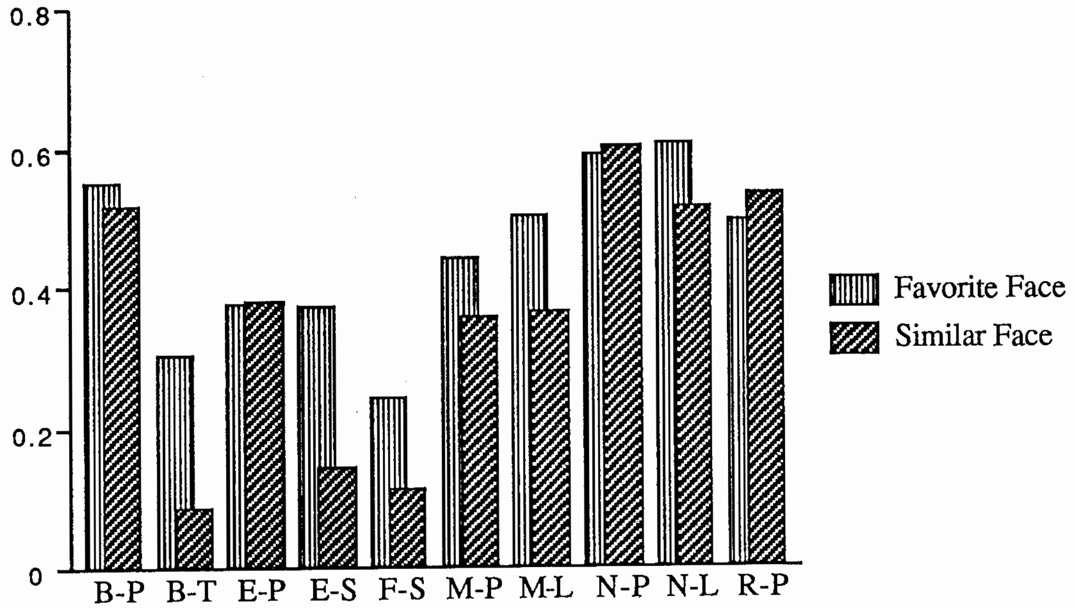

Figure 2. Variance as a function of retrieval tasks and facial features

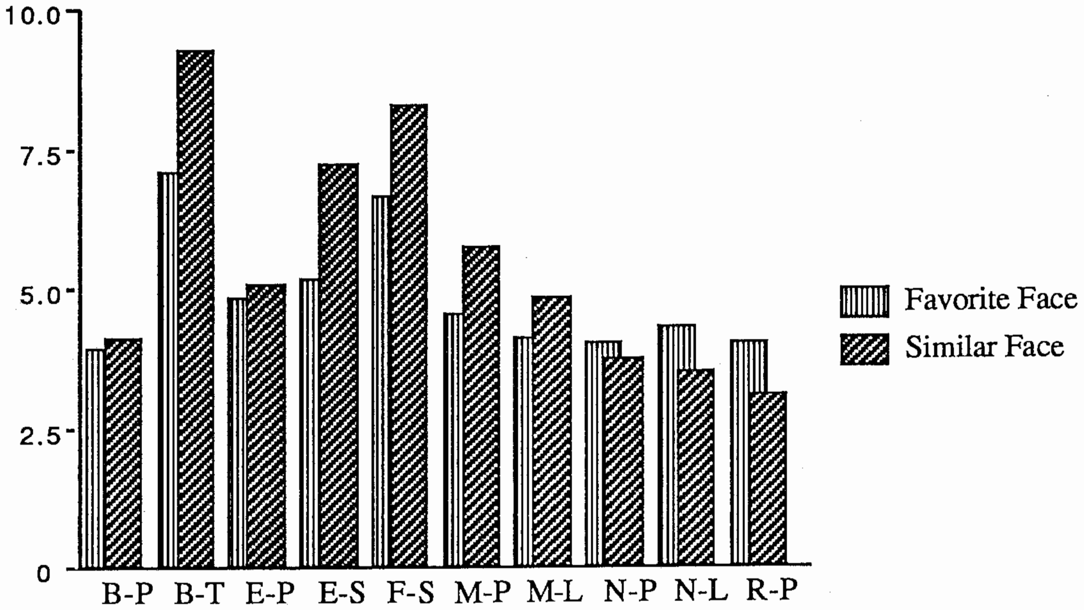

Figure 3. Hits as a function of retrieval tasks and facial features

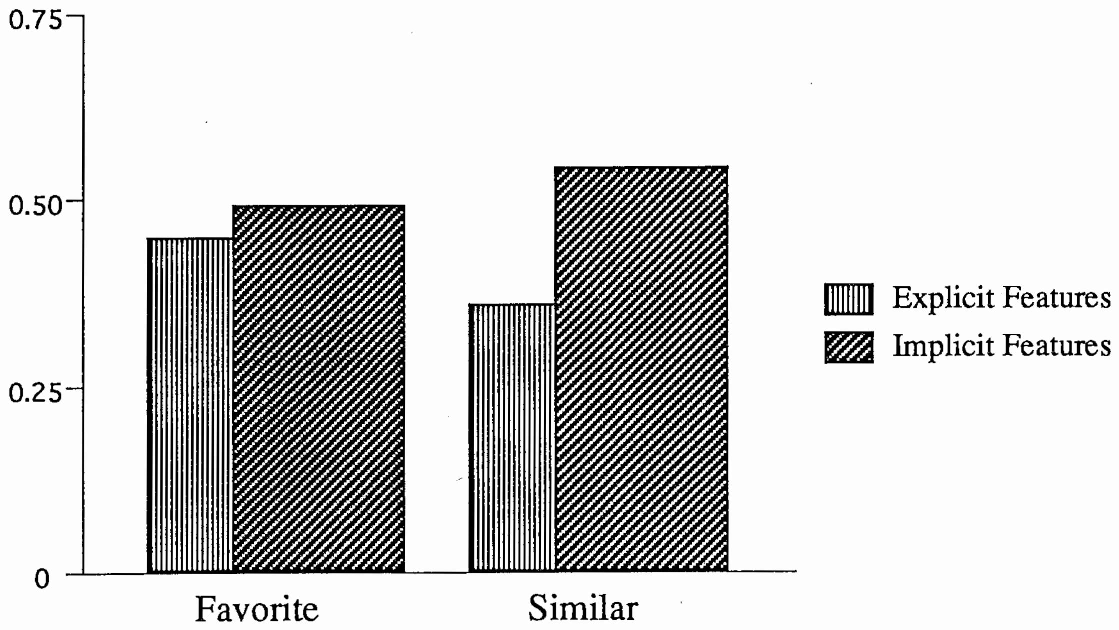

Figure 4. Variance in explicit and implicit features as a function of retrieval tasks

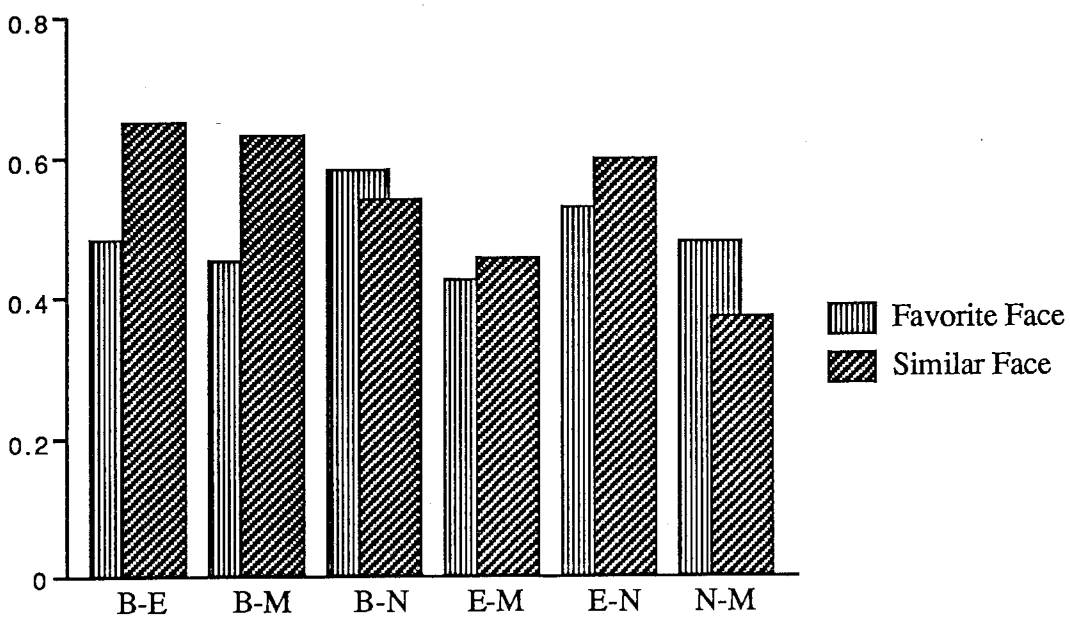

Figure 5. Variance as a function of retrieval tasks and implicit features

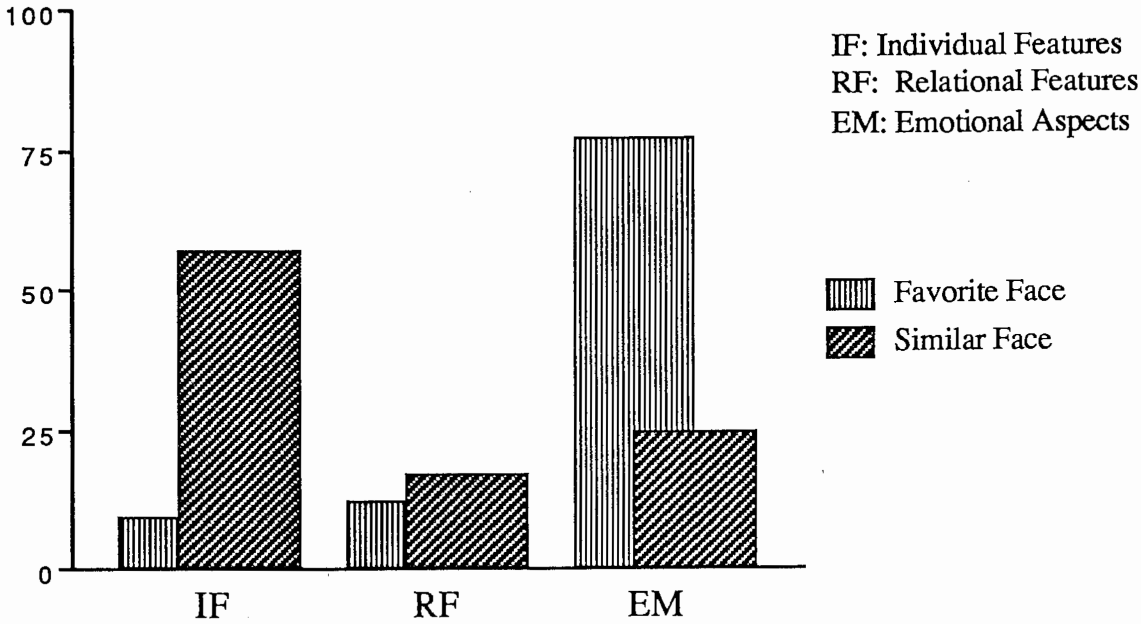

Figure 6. Reports of selection reasons as a functon of retrieval tasks