25

A Methodology for the Extraction of Reader's Emotional State Triggered from Text Typography

Dimitrios Tsonos and Georgios Kouroupetroglou
National and Kapodistrian University of Athens
Greece

1. Introduction

Writing is employed by humans in order to communicate, exchange ideas or store facts and descriptions. A well known Latin phrase is "*verba volant, scripta manent*" i.e. "spoken words fly away, written words remain". Humans in prehistoric years, used figurations to describe several events and express their fear or their admiration, the ancient Egyptians used papyrus to write down their ideas and the Chinese, in 11th Century, went beyond handwriting using typography with moveable type to create multiple copies of their documents. Through centuries typography has evolved and nowadays is one of the major manners to exchange ideas and information.

Using computers, typography and the way we write has changed radically. We have a plethora of software tools, e.g. for writing a plain text (like NotePad in MS Windows), word processors (e.g. Open Office) for text editing as well as for more elegant visual appearance of the documents and professional tools for the creation of posters, advertisements e.t.c. (e.g. Adobe's Tools and Editors). Utilising these tools, computer users have the ability not only to write a plain text, but also to format the text and arrange it in the page. In newspapers, for example, the page designer distinguishes the title from the body of text at the top of the page or the article, with larger font size. Also, when an editor wishes to emphasize a specific word or phrase he uses bold or italics typesetting. A writer can convey a message, a felling or an idea not only by the meaning of the content but also by the way the text is visually presented to the reader. Page layout affects the way a newspaper is read (Holmqvist & Wartenberg, 2005; Holmberg, 2004; Küpper, 1989; Wartenberg & Holmqvist, 2005).

The use of the WWW and the web page creation and design, introduced a new perception for the meaning of documents and publishing. In web pages, the text and background color combinations have impact on the readability and aesthetics (Porat et al., 2007; Richard & Patrick, 2004; Hill & Scharff, 1997) and a well designed graphical web document can be reader friendly (Borchers et al., 1996).

Humans express their emotions in every personal or social occasion. Everything they do or make is followed by or follows an emotional expression. For example, some people are *afraid* of being among many people. So they act accordingly by avoiding congestion. If they cannot, they are feeling *anxious* and *nervous*. Another example of the role of the emotions in

our life is that people find themselves to be more productive when they are *happy* than when being *depressed* or *unhappy*. Thus emotions play a significant role in our daily life.
But, what are emotions? There are theories that are trying to answer this question (James, 1884; Cornelius, 2000; Narayanan & Alwan, 2004) and set additional questions like which is the optimum theory to explain emotions and how can those be measured (Scherer, 2005).
This chapter aims to present how emotions can be used in Artificial Intelligence, for the automated extraction of the reader's emotional states. Starting with a survey on the basic theories on emotions, a sort description of the existing experimental procedures for measuring emotions follows. We present the Self Assessment Manikin Test (Lang, 1985) as a tool for modelling reader's emotions and emotional states such as "Pleasure", "Arousal" and "Dominance" (Tsonos et al., 2008). In the second part of the chapter there is a sort description of what typography is, what typographic elements are used in documents and how they affect the reader emotionally. In the third part we propose a system architecture for the automated extraction of readers' emotional state along with the necessary steps that should be followed during its implementation.

2. Emotions and emotional states

Through the years, several studies and theories on emotions have emerged. In this section, we describe four theories - approaches on emotions. The theoretical survey helps to understand and investigate how reader's emotions and emotional states are affected by text typography, how these emotions can be measured and the optimum experimental procedure that can be used for modelling emotions.

2.1 Theories on emotions and emotional states

Emotional theories are classified as:

- the Darwinian Theory,
- the Jamesian Theory,
- the Cognitive Theory,
- the Social Constructivist Theory.

The Darwinian Theory assumes that emotions "are evolved phenomena with important survival functions that have been selected for because they have solved certain problems we have faced as species" (Cornelius, 2000). All organisms sharing an evolutionary past should share the same emotions as well. The emotions should be analyzed in terms of their functionality and survival value (Cornelius, 2000; Narayanan & Alwan, 2004).
According to the Jamesian Theory, "bodily changes follow directly the *perception* of the exciting fact and our feeling of the same changes as they occur *is* the emotion" (James, 1884). James's belief on emotions is that, we experience emotions because our bodies have evolved to respond automatically and adaptively to features of the environment that have survival-related significance to us. Our bodies respond first and our experience of these changes constitutes what we call emotion (Cornelius, 2000). In this approach, if we consider a system that is a "black box", the stimulus (e.g. document, speech) is considered to be the input of the system while the output is the bodily change. Thus, we can change someone's emotions as we like just by changing the stimulus (Narayanan & Alwan, 2004).
The Cognitive Theory proposes that "every emotion has associated with it a particular pattern of appraisal is that if the appraisal is changed, the emotion should change as well" (Cornelius, 2000). The thought and emotions cannot be separated, and what one feels

depends in how one understands and judges the events. Emotions result from mental activity (Cornelius, 2000; Narayanan & Alwan, 2004).
Finally, according to the Social Constructivist Theory, culture and social rules play a key role in the organization of emotions at a variety of levels. The last two theories are closely related, and the development of thoughts is tightly connected to culture and social status (Narayanan & Alwan, 2004).

2.2 Experimental procedures for emotional state extraction

According to Scherer, emotion is defined as "an episode of interrelated, synchronized changes in the states of all or most of the five organismic subsystems in response to the evaluation of an external or internal stimulus event as relevant to major concerns of the organism" (Scherer, 2005).
Psychologists have developed several tools and experimental procedures that can measure emotions. They can be categorized into Free Response Measurement and Forced Choice Response Measurement.
Using Free Response Measurement, the participants of an experimental procedure are asked to respond with freely chosen labels or short expressions that characterize the nature of the emotional state they experienced. Using this procedure, instructors and researchers may face problems, like in case where the participants may not be able to express their emotions due to the use of inappropriate labels or the limited range of their vocabulary that may constrain their responses. Also, it is quite difficult to statistically analyze free responses (Scherer, 2005). Researchers sort numerous free responses into a small number of emotion categories using synonyms and resemblances. It is not a standard experimental procedure but efforts have been made towards standardizing the emotional labels.
Forced Choice Response Measurement can be subdivided (Scherer, 2005) into a) Discrete Emotion approach and b) Emotion Dimensions approach.
In Discrete Emotion approach, the participants are asked to assess their emotions using verbal expressions that best describe their emotions or provide feedback on a 3 to 5-point scale indicating whether the emotion experienced was weak or strong or use an analog scale to indicate the intensity of an experienced emotion.
There is a standardized measurement procedure for this type of experiments, but many researchers prefer to develop their own emotional categories. Such approaches result to mismatched categories that may present problems concerning the comparability of the results (Scherer, 2005). There are also problems in cross-study statistical analyses due to the abundance of missing data.
The participants in Emotion Dimensions procedure, are asked to denote how positive (pleasant) or negative (unpleasant) and how aroused (excited) or calm they feel. The emotions can be mapped using the bi-dimensional space of Pleasure-Arousal. This method is very simple, straightforward and reliable (Scherer, 2005). Also, simple or advanced single study or cross-study statistical analyses can be obtained in contrast with the Free Response Measurement and Discrete Emotion approach. In this case it is difficult to differentiate the intensity of the feeling from body excitation (Scherer, 2005).

3. Emotional state assessment using the Self-Assessment Manikin test

3.1 The S.A.M. test

The Self Assessment Manikin (SAM) Test was introduced in 1985 by P.J. Lang (Lang, 1985). The SAM procedure came up as a test for the assessment of advertisements (Bagozzi et al.,

1999). It offers the ability to avoid the verbal expression of the emotion assessment, so it establishes a quick and easy to use experimental procedure. Also, having a pictorial assessment rather than a verbal one makes SAM test cross–cultural and language-independent (Morris, 1995).

The test assesses the *emotional states* of the participants. These states are "Pleasure", "Arousal" and "Dominance" (The SAM test is also known as "PAD test" the initials of each emotional state). Synonyms are used for the expression of the PAD dimensions. "Pleasure" can be replaced by "valence" and "evaluation", "Arousal" by "activation" and "activity" and "Dominance" by "power" and "potency". In our study we used the initial verbal expressions of the three dimensions.

Using the three-dimensional space of emotional states, we are able to map them into specific emotions. A well-known example is the Russell's circumplex (Scherer, 2005). The two dimensions of "Pleasure" and "Arousal" are represented on a X-Y grid respectively. Russell placed the verbal expressions of the emotions on the grid. Figure 1 illustrates the emotional wheel with the verbal semantic mapping of the emotions.

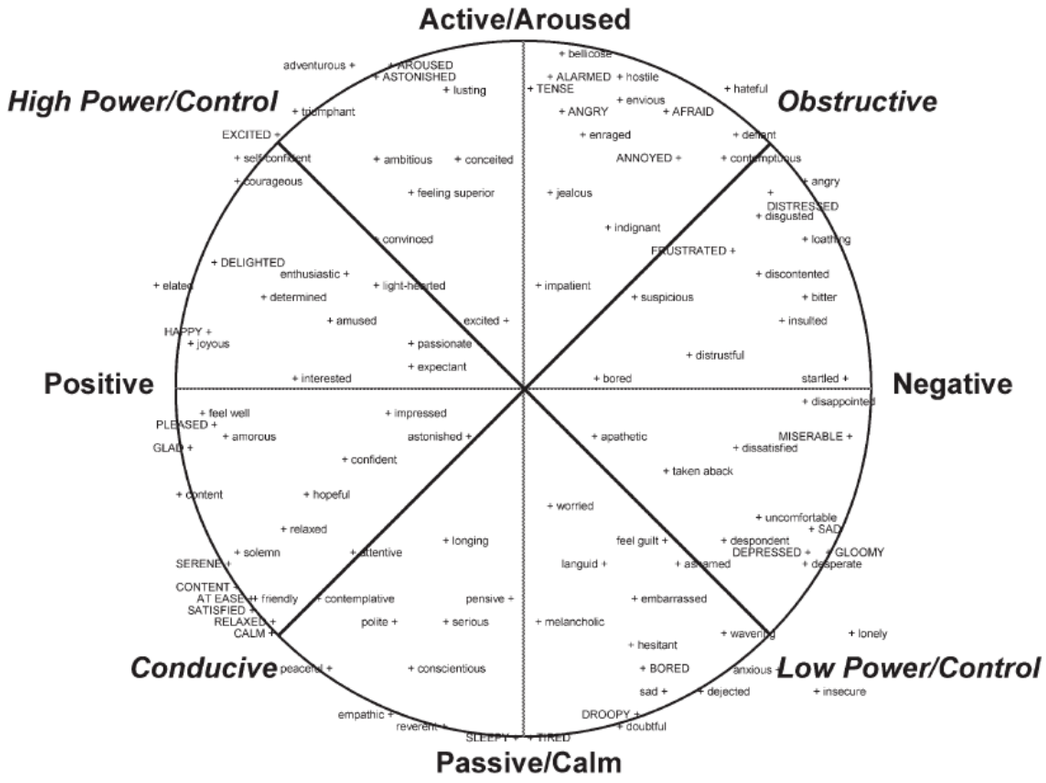


Fig. 1. The verbal expressions of the emotions placed on the Pleasure–Arousal grid according to Russell

Another version of emotional wheel is the Geneva Emotion Wheel (GEW) introduced as an experimental tool (Scherer, 2005; Bänziger et al., 2005). A computer screen of the GEW is presented in Figure 2.

How do we assess the emotional states and how do these values derive and can be presented on an X-Y grid? The SAM test, as it is described, can "produce" the values of the Pleasure, Arousal and Dominance dimensions.

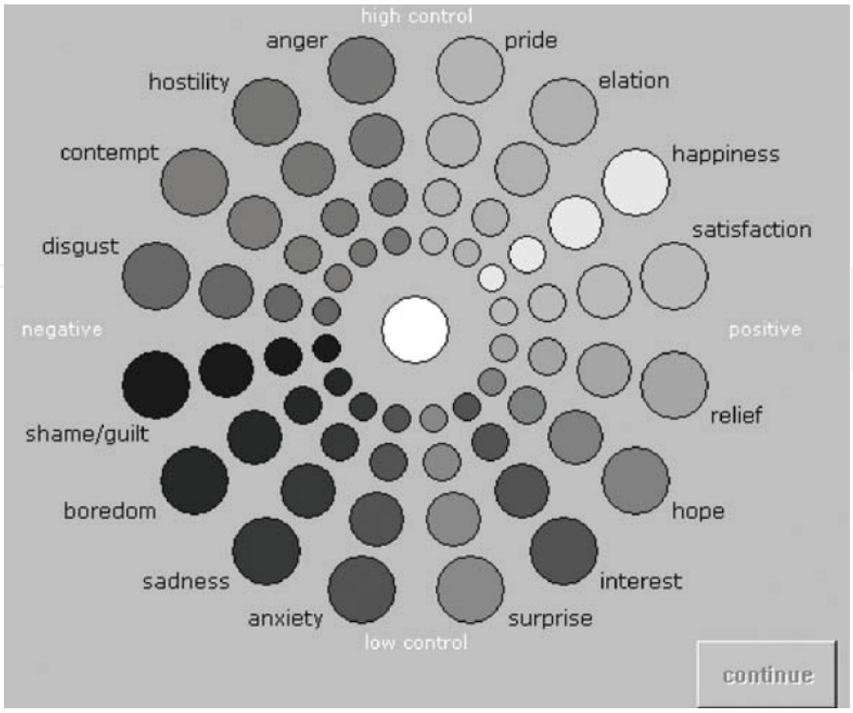


Fig. 2. A computer screen of the Geneva Emotion Wheel

During the experimental procedure, the participants assess their emotional states using five manikins (example in Figure 3) for each dimension of Pleasure, Arousal and Dominance.

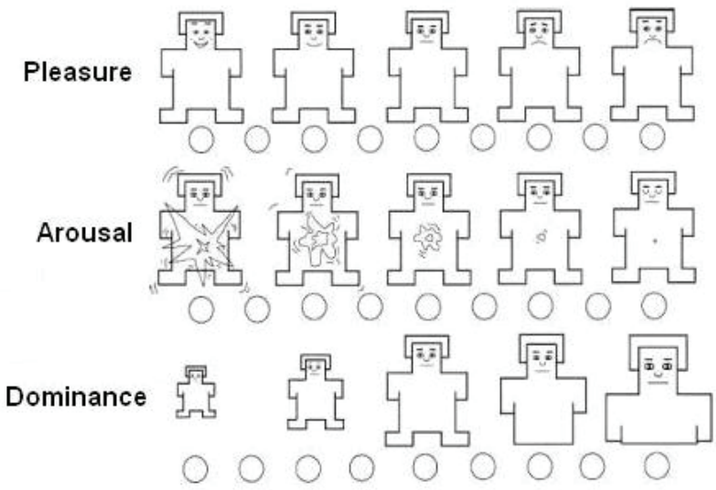


Fig. 3. The manikins of the 9-point scale SAM Test as presented during the experimental procedure. The verbal expressions of "Pleasure", "Arousal" and "Dominance" do not appear during the experimental procedure

There are no verbal expressions to assess their emotional states. The participants can choose from at least 5 selections of manikins. In some studies (Morris, 1995; Tsonos et al., 2008) are presented by larger scales of 9-points to 25-points. For example, in a 9-point scale there are 5

points of selection for the images and 4 points as interval values. The choice of the scale depends on the design of the experimental procedure and the number of the participants. If there is a need for more accuracy for the results a 9-point scale or greater should be selected. The scale of the experimental procedure is proportional to the number of subjects needed for acceptable statistical results.
For the emotional state of "Pleasure", the rating begins from a happy, smiling manikin to an unhappy, frowning one. For the "Arousal" dimension the aroused pole is represented by a highly energetic manikin and the other pole is represented by a relaxed and eyes-closed one, while for "Dominance" the controlled and in-control poles are represented by a small and large manikin respectively.
The answers of the participants can be transformed from the 1-5 point scale (or 1-9 point scale, etc.) into a dimensional space of [-1,1] or [-100%,100%]. Using the percentage approach we are able to distinguish how much an emotional state has been varied from the neutral state (the value "zero" represents the neutral state in both spaces).

3.2 Using SAM test as an experimental tool in artificial intelligence: case studies

Using SAM test we can have two different results that are closely related, the dimensional perspective of emotions and the emotions. Using only one experimental procedure, the researcher is able to have two approaches on his research.
Conducting this kind of test, modeling the emotions and emotional states of the participants, on specific projected stimuli, is feasible.
In (Grimm et al., 2006; Grimm et al., 2007a; Grimm et al., 2007b) a system was implemented for the automated estimation of emotion primitives (the dimensional approach) from speech using acoustic features.
There is also a study on the automated detection of pleasant and unpleasant emotions in spoken dialogs (Chul & Narayanan, 2005) obtained from a call center application.
Busso (Busso et al., 2007) modeled head motion sequences in expressive facial animations, analyzing them in terms of their naturalness and emotional salience in perception (SAM test is used for the evaluation of the results).

3.3 Description of the experimental procedure

In this paragraph we present a version of the SAM experimental procedure, designed for our experimental needs, following the paper and pencil IAPS Guidelines (Lang et al., 2005). The automated procedure helps to create an easy of use experiment and the rapid collection and process of the results.
For the development of the procedure, PHP was used on an Apache Web Server and MySQL. PHP allowed us to develop dynamic web pages, for the presentation of the stimuli, and to automate the registration of the participants' answers. The answers were stored in a database (MySQL). Before the experimental procedure began, participants read a short description of the purpose of our study and explicitly tutored in the emotional states theory. Certain guidelines were followed for all participants during the experimental procedure (Lang 2005). They were given ample time to read the instructions and freely ask the instructors for any additional information or clarification. Also, they were asked to complete an electronic form with some personal information, for example about their age, education level, for any visual problems, and also if she/he agrees to participate in the experimental procedure.

The participants were familiarized with the test before the actual experimental procedure. Each stimulus was presented for a few seconds after which the participant was asked to complete the form with the manikins as presented in Figure 3. By pressing the "continue" button, she/he was presented with the next stimulus and so on. The presentation of the stimuli was in the same random sequence to all the participants.

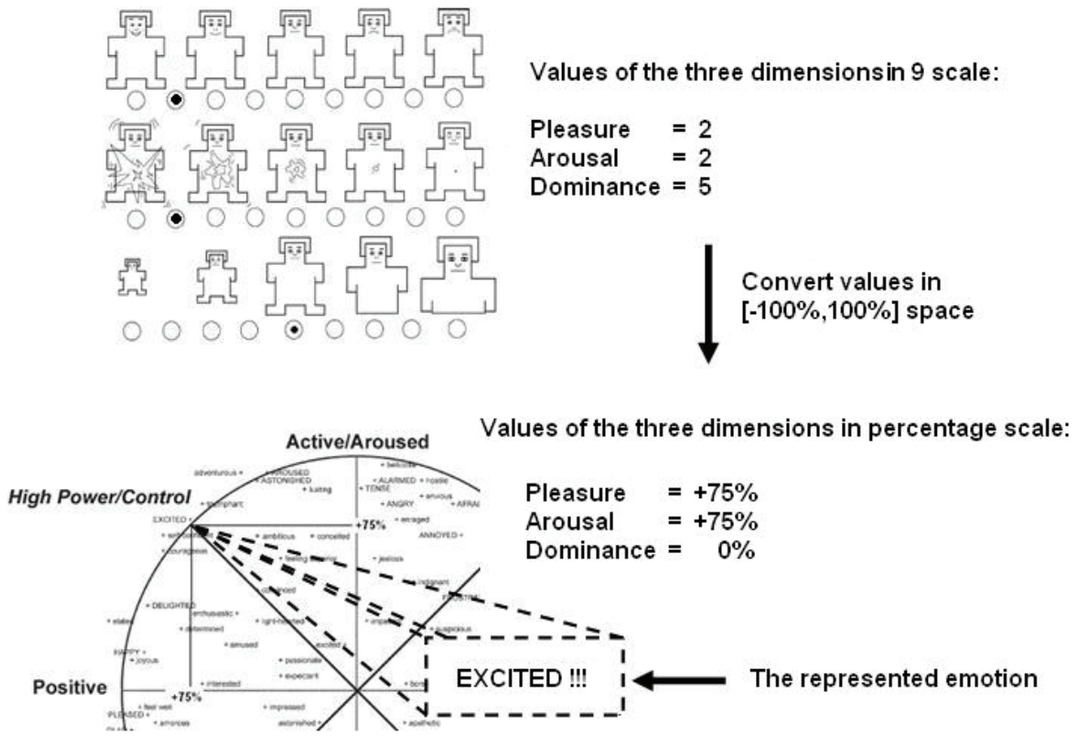


Fig. 4. Extracting the emotion using the SAM Test. Using the manikins, participants assess their emotional state. The 9-point scale values are converted into a percentage scale and then mapped on to the emotion wheel for the verbal-semantic representation of the emotion

4. Text typography in documents

The origin of the word "Typography" comes from the Greek word "Τυπογραφία" which is derived from the words "τύπος" (''τύπτω'' = to strike) and "γράφω" (=to write). Through centuries, typography has evolved using different techniques. In 20th century, Computer and Information Science created a new framework for the typography. Using computers, the layout designers of books, magazines etc. can create good aesthetic and functional results on documents in a very short time.

But in the last few years the concept "typography" has been changed due to the wide spread of computers and internet. Most computer users and readers prefer to read documents from their computer screens. The classic printed documents are transformed into electronic documents. Readers are given the possibility and functionality to interact with content of electronic documents. The reader can search specific sentences or phrases in the document, request for a summarization of specific parts of the document and browse a collection of documents.

In this section we will describe how the typographic elements can be categorized and how these elements or their combination affect the readability of the document.

4.1 Classification and usage of typographic elements

There are many efforts that categorize and classify the typographic elements, either in printed or electronic documents, according to the needs of each study. W3C (W3C, 2008a), DAISY/NISO (DAISY, 2008; ANSI/NISO, 2008) standard and Open Document Format (ODF) by (OASIS, 2008) are the three major contributors towards that classification.

In our study, the elements are categorized using the guidelines provided by these standards. The typographic elements of the documents will be mentioned as meta-data. These meta-data can be categorized in (Tsonos et al., 2007a):

- Text Formatting,
- Text Structure,
- Text Layout,
- Non-textual (Figures, Drawing, Pictures, Logos etc).

Text Formatting meta-data include the formation elements of the text, typesetting elements and font elements (like bold, italics, font size). Text Structure meta-data specify the attribute of a part of the document (chapter, title, paragraph etc.), while the Text Layout meta-data describes the visual layout of the text (like columns, headlines, borders). The text with the formation and structural metadata can be combined with other non-textual metadata, such as figures, drawing etc.

One can notice that there is a relation between these elements. For example, a title (text structure) element may have a 16pt font size (text formatting) and placed in a text column (text layout). A subtitle element may have 14pt font size and placed in a column, but also under the title.

Often, textual features present purely logical structure information. For example, a piece of text can be part of a list or footnote. However, frequently, the use of these features has the purpose of communicating semantic information to the reader. Examples of such purposes are: giving emphasis to an important piece of text; focusing the readers' attention to the central point of an article; highlighting a name entity like the title of a movie and so on. This classification can help the creation of sophisticated document manipulation tools including an enhanced audio representation of the document.

Recently there was an attempt to produce an automatic extraction system of semantic information based only on the document layout, without the use of natural language processing (Fourli-Kartsouni et al., 2007). However, there are several studies on the automatic identification of logical structure of documents e.g. (Conway, 1993; Yamashita et al., 1991; Derrien-Peden, 1991; Krishnamoorthy, 1993). Most traditional approaches in this field have employed deterministic methods (decision trees, formal grammars) (Mao et al., 2003; Tsujimoto & Asada, 1990; Derrien-Peden, 1991), which may suffer from poor performance due to noise and uncertainty. In addition, such approaches create models which are not flexible to domain changes and cannot easily evolve in the presence of new evidence.

In order to overcome such limitations, Fourli-Kartsouni (Fourli-Kartsouni et al., 2007) employed a probabilistic approach based on Bayesian networks trained on a series of labelled documents. Bayesian networks offer a significant tolerance to noise and uncertainty, they can capture the underlying class structure residing in the data and they can be trained on examples, thus adapting to existing and future evidence.

4.2 How typographic elements affect reader's emotional state, comprehension and navigation in electronic documents?

E-documents have exceeded documents in printed format by providing several functionalities to the reader like browsing, navigation, searching, highlighting and multimedia facilities. The meta-data and the way it is combined, has specific meaning for each reader in different documents (Küpper, 1989; Holmberg, 2004) and the way the document is read (Holmqvist et al., 2003; Holmqvist & Wartenberg, 2005). The authors utilize structural and layout meta-data to personalize the content. The style of a document sets the typographic rules, used more often in technical or scientific papers and reports than in magazines. There are also studies on how the combination of colors on a page and the different type of style on text, affects the readers' emotional state and the readability of the document (Laarni, 2003), the meaning (Küpper, 1989; Holmberg, 2004) and the reader comprehension.

Emotions and the emotional state of the reader depend on document structure, layout and text formatting. Multiple combinations of colors (Birren, 1984), font size, type and style in a document affects the emotional state (Sánchez et al., 2006; Sánchez et al., 2005) and consequently the readability of the document (Laarni, 2003; Saari et al., 2004) not only in printed but also in electronic format (Larson, 2007).

5. Extraction and modeling reader's emotional state

5.1 Proposed architecture

Based on the SAM test, we proposed an XML-based architecture for the real-time extraction of reader's emotional state *excluding* any content and/or domain dependent information from the input documents (Tsonos et al., 2007b). Figure 5 illustrates a diagram of the system architecture. The proposed system is designed to process all types of documents in printed or electronic format. The Markup Normalization Module converts all non-tagged documents as well as tagged (not conforming to the DAISY/NISO format) into tagged compliant to DAISY/NISO format. Printed documents are scanned and parsed through an OCR system so to be digitized and exported in a tagged format. Documents being already in a tagged format include meta-data about the format and the structure of the text. All these have to be normalized to the required meta-data and file type (DAISY/NISO standard).

The documents, in the desired format, can be processed by the Emotional State Extraction Module. This module is implemented using the model derived by several experimental procedures on how the reader is affected by the document metadata provided in paragraph 4.1. The experimental procedure is similar to one described in 3.3. Conducting multiple experiments using SAM Test we are able to distinguish the way each document element affects the reader's emotional state but also their combination. For the readers' emotional state modelling in paragraph 5.3, the results from ongoing experiments and how the readers' emotional states varying according to font type, size and color, background color and typesetting elements are presented. Future work may reveal the need of a rule based model or a statistical one.

The Emotional State Extraction Module produces an XML file, the Emotional-ML. This file contains the content of the documents but also the initial tagging (after markup normalization) and the emotional state annotation. Currently, there is an effort to create a new markup language by W3C Emotion Markup Language Incubator Group. The group will discuss and propose scientifically valid representations of several aspects of emotional states that appear to be relevant for a number of use cases (W3C, 2008b).

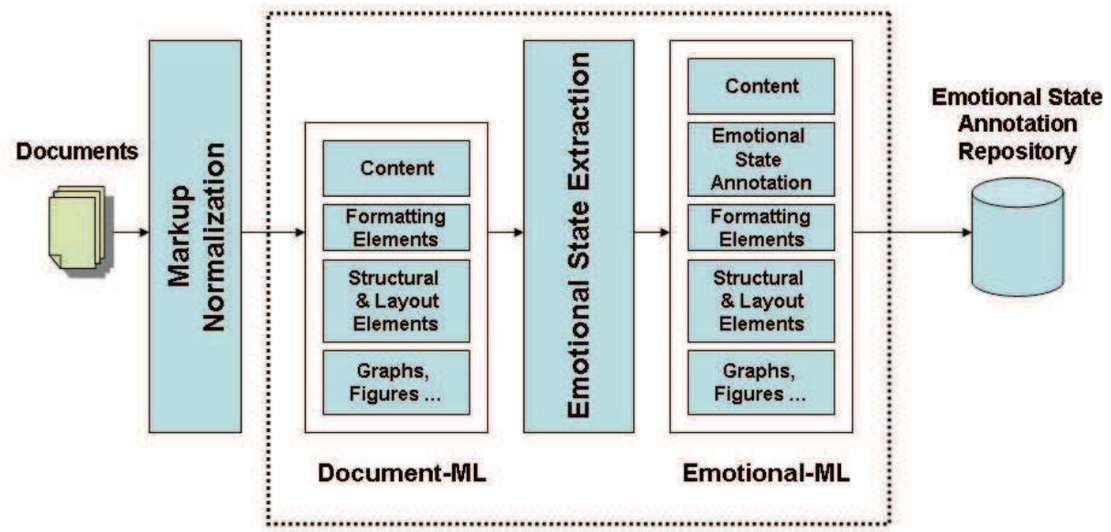


Fig. 5. The proposed architecture for the real-time extraction of reader's emotional state

5.2 Experimental methodology

For the reader's emotional state modelling, an experimental procedure has been created according to 3.3, assessing 15 participants. Fifty-four combinations of font and typesetting elements, projected on identical LCD displays in 1024x768 resolution, were investigated on an emotionally neutral Greek text similar to one used by Hill & Scharff (Hill & Scharff, 1997). These are:

- Plain text in font sizes 9pt, 10pt, 11pt, 12pt, 13pt, 14pt, 15pt, 16pt, 18pt, 26pt and 32pt and font type Times New Roman.
- Plain, bold, italics, bold-italics in size 16pt. in both Arial and Times New Roman, in color combinations as proposed in (Laarni, 2003) and (Hill & Scharff, 1997) (see Table 1).

During the experimental procedure, the presentation of some stimuli was repeated for two or three times, in order to correlate the dependency of the previously presented combinations of typographic elements to the current reader's emotional response. All stimuli were displayed in a random sequence for approximately 15 seconds. In the following screen participants were asked to assess their emotional state on a 9-point PAD scale using the manikins provided by the SAM test.

The duration of the experimental procedure was 20-30 minutes, depending on the participants response time.

5.3 Results

The results indicated that during the experimental procedure, the participants were mostly influenced by the font/background color combinations and the font size rather than by the typesetting elements and font type. For the later, a new experimental procedure should be designed and implemented that will clearly distinguish the affect of these elements to the emotional states.

Previously studied (Laarni, 2003) font and background color combinations as stimuli in the experimental procedure, was to investigate the hypothesis and the confirmation of the results that the SAM test is cross-cultural and language independent (Morris, 1995) test.

The mean values of the emotional states are displayed in Figure 6. According to (Laarni, 2003) and the presented experiment, the outcome can be summarized as:

- Red on Green (RG) color combinations are the most arousing and less pleasant.
- Black on White (BW) color combination has the lowest mean arousal value.
- White on Blue (WU), Green on Yellow (GY) and Black on White (BW) are the most pleasant combinations.

Abbreviation	Font Color	Background Color
YU	Yellow	Blue
WB	White	Black
GY	Green	Yellow
BW	Black	White
BG	Black	Grey
RG	Red	Green
WU	White	Blue

Table 1. The explaination of the abbreviations used in Fig. 6

The font size has a great impact to the reader's emotional state and especially to "Pleasure" and "Dominance" dimensions as presented in Figure 7. The "small-size fonts" (9 to 13 points) can be characterized as unpleasant and aroused. The "medium-size fonts" (14 to 27 points) are considered to be the most pleasant and calm. The "large-size fonts" (27 points and higher) have similar impact with "small size fonts".

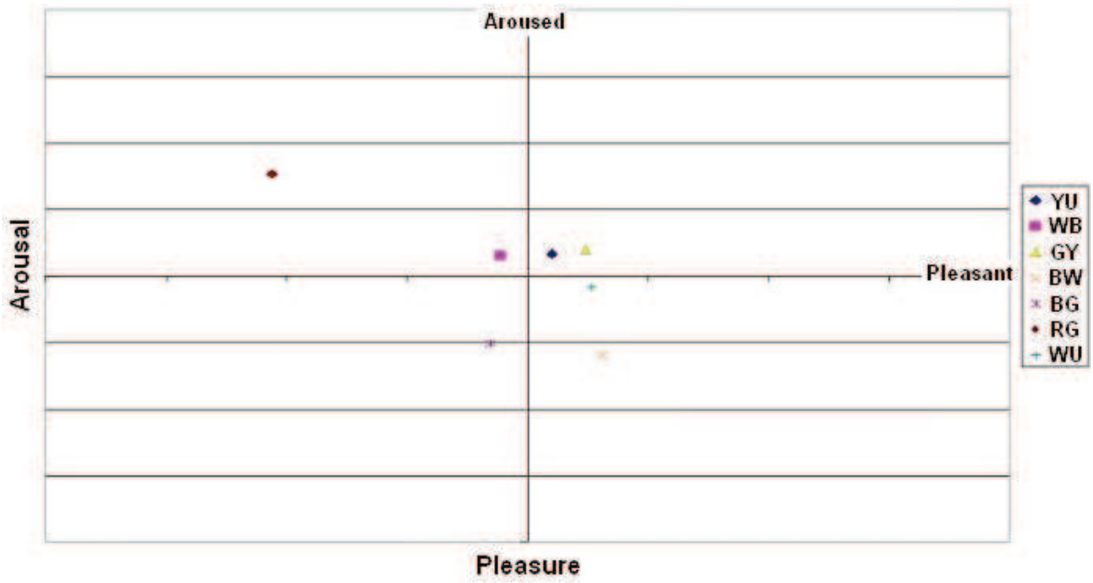


Fig. 6. The font and background color combinations on "Pleasure" and "Arousal" grid. The abbreviation of the color combinations are explained in Table 1

The statistical analysis of our results, is based on the approach used by (Grimm & Kroschel, 2005) on listeners' emotional state assessment. From equations (1), (2), (3) and (4), the resulting mean values of the participants-evaluators confidence scores ($\bar{r}_k$) are shown in Table 2.

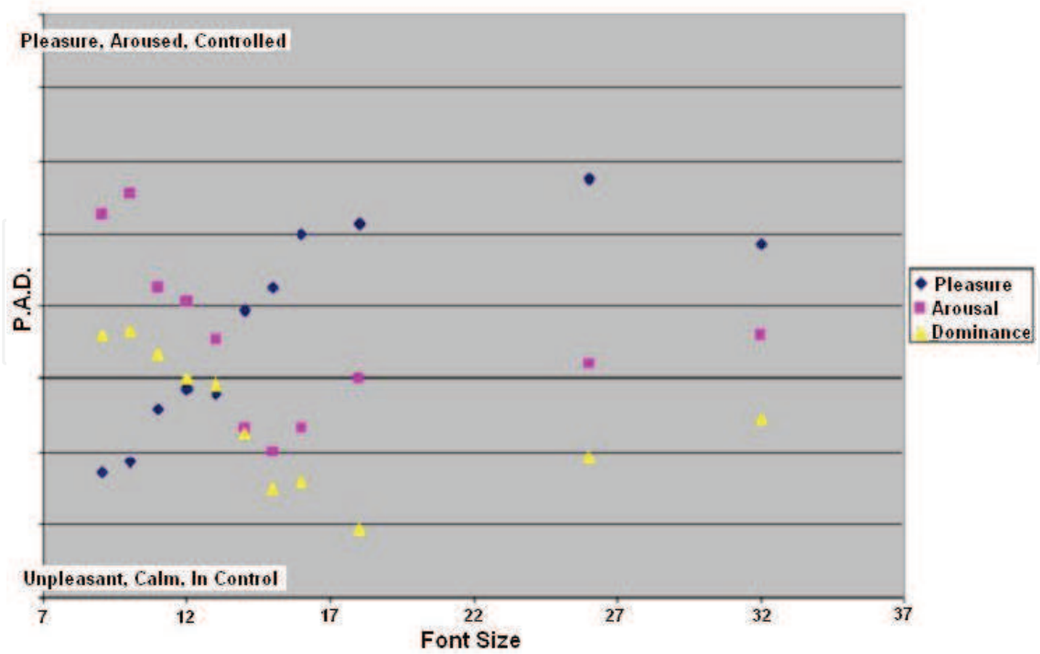


Fig. 7. The distribution of the impact of font size on the 3 dimensions of "Pleasure", "Arousal" and "Dominance"

	Mean Values			
Number of evaluators	Age	$\bar{r}_k^P$ for "Pleasure"	$\bar{r}_k^A$ for "Arousal"	$\bar{r}_k^D$ for "Dominance"
15	26,3	0,59	0,51	0,53

Table 2. The participants mean confidence score for each dimension

$$x_n^{MLE,(i)} = \frac{1}{K}\sum_{k=1}^{K} x_{n,k}^{(i)} \tag{1}$$

$$\mu_k^{(i)} = \frac{1}{N}\sum_{n=1}^{N} x_{n,k}^{(i)} \tag{2}$$

$$\mu^{MLE,(i)} = \frac{1}{N}\sum_{n=1}^{N} x_n^{MLE,(i)} \tag{3}$$

$$r_k = \frac{\sum_{n=1}^{N}(x_{n,k}^{(i)} - \mu_k^{(i)})(x_n^{MLE,(i)} - \mu^{MLE,(i)})}{\sqrt{\sum_{n=1}^{N}(x_{n,k}^{(i)} - \mu_k^{(i)})^2}\sqrt{\sum_{n=1}^{N}(x_n^{MLE,(i)} - \mu^{MLE,(i)})^2}} \tag{4}$$

An ideal evaluator is considered with $r_k^{(i)} = 1$ and an unreliable one with $r_k^{(i)} < 0$. All evaluators are reliable ($r_k^{(i)} > 0$).

6. Conclusion, future work and potential applications

In this chapter we presented a theoretical survey on emotions and a methodology on how reader's emotional state can be modelled. The selection of the Self Assessment Manikin Test and the dimensional approach for emotions led to a procedure for modelling the reader's emotional states on typesetting and font elements of e-documents. Documents contain many more elements and their combinations as presented in 4.1 prior to the elements in our experiment, for further experimentation.

The proposed modelling can be used in Artificial Intelligence for a number of applications. For example, as proposed in (Tsonos et al., 2007b), we can use the emotionally annotated documents for the multimodal accessibility of documents, focusing on the acoustic modality. Expressive Speech Synthesis (IEEE, 2006) creates more physical results. The dimensional approach of the emotions is also used in speech synthesis (Tsonos et al., 2007b) (Schröder, 2006). In Figure 8 a proposed real-time system that automatically produces emotional annotation to documents and conveys the visual elements into acoustic modality using expressive speech synthesis is presented.

Future work includes the use of an e-TSA composer (Xydas & Kouroupetroglou, 2001a, 2001b) (Xydas et al., 2005) on the DEMOSTHeNES Text-to-Speech platform (Xydas & Kouroupetroglou, 2001c) and models for expressive speech synthesis as proposed by Schröder (Schröder, 2006), for acoustic rendition of emotionally annotated documents.

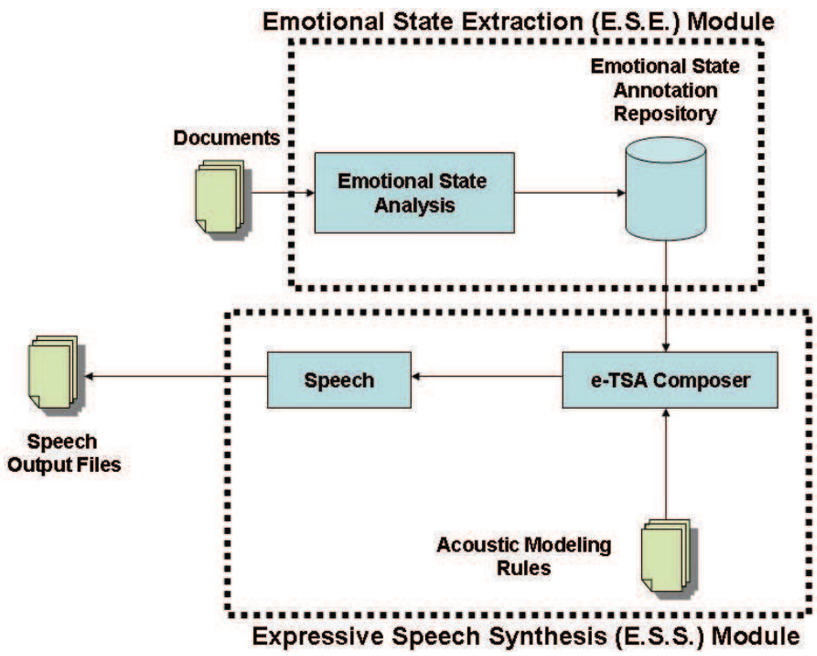


Fig. 8. The proposed architecture produces automatically emotional state annotation and uses Expressive Speech Synthesis for the acoustic production of documents

7. Acknowledgements

The work described in this chapter has been funded by the European Social Fund and Hellenic National Resources under the HOMER project of the Programme: PENED, Greek General Secretariat of Research and Technology.

8. References

ANSI/NISO (2008), Specifications for the digital talking book, *http://www.niso.org/kst /reports/standards/*

Bagozzi, R. P.; Gopinath, M. & Nyer, P. U. (1999). The Role of Emotions in Marketing. *Journal of the Academy of Marketing Science*, Vol. 27, No. 2, pp. 184 – 206, ISSN 1552-7824

Bänziger, T.; Tran, V., & Scherer, K. (2005). The Geneva Emotion Wheel: a tool for the verbal report of emotional reactions, *Poster presented at ISRE 2005*, Bari, Italy, (Available as pdf from http://www.unige.ch/fapse/emotion/isre/ BanzigerTranSchererISRE05.pdf)

Birren, F. (1984). *Color & Human Response: Aspects of Light and Color Bearing on the Reactions of Living Things and the Welfare of Human Beings*. J. Wiley & Sons Inc, ISBN: 978-0-471-28864-0, New York

Borchers, J.; Deussen, O.; Klingert, A. & Knörzer, C. (1996). Layout rules for graphical Web documents. *Computers and Graphics*, Vol. 20, No. 3, pp. 415 – 426, ISSN: 0097-8493

Busso, C.; Deng, Z.; Grimm, M.; Neumann, U. & Narayanan, S. (2007). Rigid Head Motion in Expressive Speech Animation: Analysis and Synthesis. *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 15, No. 3, March 2007, pp. 1075-1086, ISSN: 1558-7916

Chul, M. L. & Narayanan, S. S. (2005). Toward detecting emotions in spoken dialogs. *IEEE Transactions on Speech and Audio Processing*, Vol. 13, No. 2, March 2005, pp. 293-303, ISSN: 1063-6676

Cornelius, R. R. (2000). Theoretical approaches to emotion, *Proceedings of ISCA Workshop on Speech and Emotion*, pp. 3–10, Belfast, Northen Ireland, September 2000

Conway, A. (1993). Page grammars and page parsing: a syntactic approach to document layout recognition, *Proceedings of the Second International Conference on Document Analysis and Recognition* 1993, pp. 761-764, ISBN: 0-8186-4960-7, Tsukuba Science City, Japan, 20 – 22 October 1993

DAISY Consortium (2008), *www.daisy.org*

Derrien-Peden, D. (1991). Frame-based system for macro-typographical structure analysis in scientific papers. *Proceedings of International Conference on Document Analysis and Recognition*, pp. 311-319, Saint-Malo, France, September 1991

Grimm, M. & Kroschel, K. (2005). Evaluation of natural emotions using self assessment manikins, *Proceedings of the IEEE Workshop on Automatic Speech Recognition and Understanding*, pp. 381-385, ISSN: 0-7803-9478-x, 27 Nov. – 1 Dec. 2005

Grimm, M.; Mower, E.; Narayanan S. & Kroschel, K. (2006). Combining categorical and primitives-based emotion recognition, *Proceedings of 14th European Signal Processing Conference (EUSIPCO)*, Florence, Italy, September 2006

Grimm, M.; Kroschel, K.; Mower, E. & Narayanan S. (2007a). Primitives-based evaluation and estimation of emotions in speech. *Speech Communication*, Vol. 49, No. 10-11, October 2007, pp. 787-800, ISSN: 0167-6393

Grimm, M.; Kroschel, K. & Narayanan, S.(2007b). Support Vector Regression for Automatic Recognition of Spontaneous Emotions in Speech, *IEEE International Conference on Acoustics, Speech and Signal Processing 2007 (ICASSP 2007)*, Vol. 4, pp. 1085-1088, ISBN: 1-4244-0728-1, Hawaii, USA, 15 – 20 April 2007

Hill, A. & Scharff, L. V. (1997). Readability of screen displays with various foreground/background color combinations, font styles, and font types, *Proceedings of the Eleventh National Conference on Undergraduate Research (NCUR-97)*, Vol. 2, pp. 742-746, Austin, Texas U.S.A., 24-26 April 1997

Holmberg, N. (2004). *Eye movement patterns and newspaper design factors. An experimental approach*. Master Thesis, Lund University Cognitive Science

Holmqvist, K.; Holsanova, J.; Barthelson, M. & Lundqvist, D. (2003). Reading or scanning? A study of newspaper and net paper reading, In: *The mind's eye: cognitive and applied aspects of eye movement research*, Hyönä, J. R., and Deubel, H. (Eds.), pp. 657-670, Elsevier Science Ltd, ISBN: 0-444-51020-6, Holland

Holmqvist, K. & Wartenberg, C. (2005). The role of local design factors for newspaper reading behaviour – an eye-tracking repspective, *Lund University Cognitive Studies*, 127. Lund: LUCS, ISSN: 1101-8453

IEEE (2006). Special Section on Expressive Speech Synthesis, *as presented in the Editorial of IEEE Transactions on Audio, Speech and Language Processing*, Vol. 14, No 4, July 2006, pp. 1097-1098, ISSN: 1558-7916

James, W. (1884). What is an emotion? *Mind*, Vol. 19, pp. 188 – 205, ISSN: 0026-4423

Fourli-Kartsouni, F. ; Slavakis, K. ; Kouroupetroglou, G. & Theodoridis S. (2007). A Bayesian Network Approach to Semantic Labelling of Text Formatting in XML Corpora of Documents, *Lecture Notes in Computer Science (LNCS)* Vol. 4556, pp. 299-308, ISBN : 978-3-540-73282-2

Krishnamoorthy, M.; Nagy, G.; Seth, S. & Viswanathan, M. (1993). Syntactic segmentation and labeling of digitized pages from technical journals. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, Iss. 7, July 1993, pp. 737-747, ISSN: 0162-8828

Küpper, N. (1989). Recording of Visual Reading Activity: Research into Newspaper Reading Behaviour, (Available as pdf from *http://calendardesign.de/leseforschung/ eyetrackstudy.pdf*)

Laarni, J. (2003). Effects of color, font type and font style on user preferences, *Adjunct Proceedings of HCI International 2003*, C. Stephanidis (Ed.), Greece, pp. 31-32, Heraklion, University Press, ISBN: 960-524-166-8

Lang, P. J. (1985). The cognitive psychophysiology of emotion: Fear and anxiety, In: *Anxiety and the anxiety disorder*, A. Tuma & J. Maser (Eds.), pp. 131-170, Lawrence Erlbaum, New Jersey

Lang, P. J.; Bradley, M. M. & Culthbert, B. N. (2005). *International Affective Picture System (IAPS): Instruction Manual and Affective Ratings*. Technical Report A-6, The Center for Research in Psychophysiology, University of Florida, U.S.A.

Larson, K. (2007). The Technology of Text. *IEEE Spectrum*, May 2007, pp. 20 – 25, ISSN: 0018-9235

Mao, S; Rosenfeld, A. & Kanungo, T. (2003). Document structure analysis algorithms: a literature survey, *Proceedings of SPIE*, Vol. 5010, pp. 197-207, ISBN 0-8194-4810-9

Morris, J. D. (1995). Observations SAM: The self-assessment manikin- An efficient cross-cultural measurement of emotional response, *Journal of Advertising Research*, Vol. 6, November - December 1995, pp. 63-68, ISSN: 1470-7853

Narayanan, S. & Alwan, A. (2004). *Text to Speech Synthesis : new paradigms and advances*, Prentice Hall, ISBN: 9-780131-456617, Upper Sandle River, New Jersey

OASIS (2008). Organization for the Advancement of Structured Information Standards, *www.oasis-open.org/home/index.php*

Porat, T.; Liss, R. & Tractinsky, N. (2007). E-Stores Design: The Influence of E-Store Design and Product Type on Consumers' Emotions and Attitudes. *Lecture Notes in Computer Science (LNCS)*, Vol. 4553, pp. 712 - 721, ISBN: 978-3-540-73109-2

Richard, H. & Patrick, H. (2004). The Impact of Web Page Text-Background Colour Combinations on Readability, Retention, Aesthetics and Behavioural Intention. *Behaviour and Information Technology*, Vol. 23, No. 3, May-June 2004, pp.183-195, Taylor & Francis Group Journals, ISSN-0144-929X

Saari, T.; Turpeinen, M.; Laarni, J.; Ravaja N. & Kallinen, K. (2004). Emotionally Loaded Mobile Multimedia Messaging, *Third International Conference Entertainment Computing - ICEC 2004*, pp. 476-486, Eindhoven, The Netherlands, 1-3 September 2004

Sánchez, J. A.; Kirschning, I.; Palacio, J. C. & Ostróvskaya, Y. (2005). Towards mood-oriented interfaces for synchronous interaction, *Congreso Latinoamericano de Interacción Humano-Computadora (CLIHC´05)*, pp. 1-7, Cuernavaca, México, 2005

Sánchez, J. A.; Hernández, N. P.; Penagos J. C. & Ostróvskaya, Y. (2006). Conveying mood and emotion in instant messaging by using a two-dimensional model for affective states, *Proceedings of the Symposium on Human Factors in Computer Systems (IHC 2006)*, pp. 66-72, ISBN: 1-59593-432-4, Brazil, 2006, ACM, New York

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, Vol. 44, No. 4, pp. 693-727, ISSN: 0539-0184

Schröder, M. (2006). Expressing degree of activation in synthetic speech. *IEEE Transactions on Audio, Speech and Language Processing*, Vol. 14, Iss. 4, July 2006, pp. 1128-1136, ISSN: 1558-7916

Tsonos, D.; Xydas, G. & Kouroupetroglou G. (2007a). Auditory Accessibility of Metadata in Books: A Design for All Approach. *Lecture Notes in Computer Science (LNCS)*, Vol. 4556, pp. 436 - 445, ISBN: 978-3-540-73282-2

Tsonos, D.; Xydas, G. & Kouroupetroglou, G. (2007b). A Methodology for Reader's Emotional State Extraction to Augment Expressions in Speech Synthesis, *Proceedings of 19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007)*, Vol. 2, pp. 218-225, Patras, Greece, 29-31 October 2007 ISBN: 0-7695-3015-X

Tsonos, D.; Ikospentaki, K. & Kouroupetrglou, G. (2008). Towards Modeling of Readers' Emotional State Response for the Automated Annotation of Documents, *IEEE World Congress on Computational Intelligence (WCCI 2008)*, pp. 3252 - 3259, Hong Kong, 1 – 6 June 2008, ISSN: 978-1-4244-1821-3

Tsujimoto, S. & Asada, H. (1990). Understanding multi-articled document, *Proceedings of 10th International Conference on Pattern Recognition*, Vol. I, pp. 551-556, ISBN: 0-8186-2062-5, Atlantic City, NJ, USA, 16-21 June 1990

Wartenberg, C. & Holmqvist, K. (2005). Daily Newspaper Layout - Designers' Predictions of Readers' Visual Behaviour - A Case Study, *Lund University Cognitive Studies*, 126. Lund: LUCS, ISSN: 1101-8453

W3C (2008a), World Wide Web Consortium, *www.w3.org*

W3C (2008b), World Wide Web Consortium, Emotion Markup Language Incubator Group, *www.w3.org/2005/Incubator/emotion*

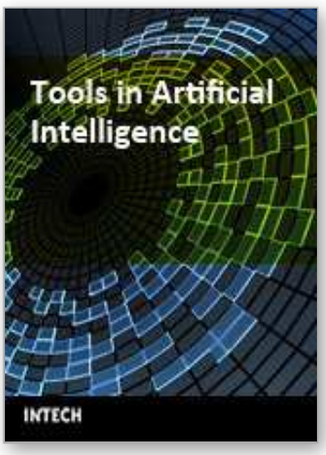
Xydas, G.; Argyropoulos, V.; Karakosta, T. & Kouroupetroglou, G. (2005). An Experimental Approach in Recognizing Synthesized Auditory Components in a Non-Visual Interaction with Documents, *Proceedings of the 11th Int. Conference on Human-Computer Interaction*, Las Vegas, Nevada USA, 22-27 July 2005, Lawrence Erlbaum Associates, Inc, ISBN: 0-8058-5807-5

Xydas G. & Kouroupetroglou, G. (2001a). Text-to-Speech Scripting Interface for Appropriate Vocalisation of e-Texts, *Proceedings of EUROSPEECH 2001*, pp. 2247-2250, Aalborg, Denmark, 2-7 September 2001, ISBN: 87-90834-09-7

Xydas, G. & Kouroupetroglou, G. (2001b). Augmented Auditory Representation of e-Texts for Text-to-Speech Systems, *Lecture Notes in Artificial Intelligence (LNAI)*, Vol. 2166, 2001, pp. 134-141, ISBN: 978-3-540-42557-1

Xydas, G. & Kouroupetroglou, G. (2001c). The DEMOSTHeNES Speech Composer, *Proceedings of the 4th ISCA Tutorial and Research Workshop on Speech Synthesis*, pp. 167-172, Perthshire, Scotland, 29 August – 1 September 2001

Yamashita, A.; Amano, T.; Takahashi, I. & Toyokawa, K. (1991). A model based layout understanding method for the document recognition system, *Proceedings of International Conference on Document Analysis and Recognition*, pp. 130-138, Saint Malo, France, September 1991



Tools in Artificial Intelligence

Edited by Paula Fritzsche

ISBN 978-953-7619-03-9
Hard cover, 488 pages
Publisher InTech
Published online 01, August, 2008
Published in print edition August, 2008

This book offers in 27 chapters a collection of all the technical aspects of specifying, developing, and evaluating the theoretical underpinnings and applied mechanisms of AI tools. Topics covered include neural networks, fuzzy controls, decision trees, rule-based systems, data mining, genetic algorithm and agent systems, among many others. The goal of this book is to show some potential applications and give a partial picture of the current state-of-the-art of AI. Also, it is useful to inspire some future research ideas by identifying potential research directions. It is dedicated to students, researchers and practitioners in this area or in related fields.

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen