Article

Investigation of Relationships between Discrete and Dimensional Emotion Models in Affective Picture Databases Using Unsupervised Machine Learning

Marko Horvat [1,*], Alan Jović [2] and Kristijan Burnik [3]

1 Department of Applied Computing, Faculty of Electrical Engineering and Computing, University of Zagreb, Unska 3, HR-10000 Zagreb, Croatia
2 Department of Electronics, Microelectronics, Computer and Intelligent Systems, Faculty of Electrical Engineering and Computing, University of Zagreb, Unska 3, HR-10000 Zagreb, Croatia
3 Independent Researcher, HR-10000 Zagreb, Croatia
* Correspondence: marko.horvat3@fer.hr

Featured Application: We have shown that affective ratings in picture stimuli are correlated: (1) there is a statistically significant correlation between specific pairs of discrete and dimensional emotions in picture stimuli, and (2) robust transformation of picture ratings from the discrete emotion space to well-defined clusters in the dimensional space is possible for some discrete-dimensional emotion pairs. These findings can be used to improve the retrieval of stimuli from affective repositories. Based on these results, an efficient system for recommending affective multimedia can be developed.

Abstract: Digital documents created to evoke emotional responses are intentionally stored in special affective multimedia databases, along with metadata describing their semantics and emotional content. These databases are routinely used in multidisciplinary research on emotion, attention, and related phenomena. Affective dimensions and emotion norms are the most common emotion data models in the field of affective computing, but they are considered separable and not interchangeable. The goal of this study was to determine whether it is possible to statistically infer values of emotionally annotated pictures using the discrete emotion model when the values of the dimensional model are available and vice versa. A positive answer would greatly facilitate stimuli retrieval from affective multimedia databases and the integration of heterogeneous and differently structured affective data sources. In the experiment, we built a statistical model to describe dependencies between discrete and dimensional ratings using the affective picture databases NAPS and NAPS BE with standardized annotations for 1356 and 510 pictures, respectively. Our results show the following: (1) there is a statistically significant correlation between certain pairs of discrete and dimensional emotions in picture stimuli, and (2) robust transformation of picture ratings from the discrete emotion space to well-defined clusters in the dimensional space is possible for some discrete-dimensional emotion pairs. Based on our findings, we conclude that a feasible recommender system for affective dataset retrieval can be developed. The software tool developed for the experiment and the results are freely available for scientific and non-commercial purposes.

Keywords: picture stimuli; affective pictures databases; statistical analysis; clustering; emotion; affective computing

1. Introduction

Digital media, apart from having a format, semantic content, and context, also evoke emotions. People who watch movies on online video streaming services, listen to music on digital media, read books, or use virtual reality are intrinsically emotionally stimulated.

This process of eliciting emotional responses in exposed individuals is multifaceted, spontaneous, and has complex neuroanatomical underpinnings [1,2]. The specific nature and magnitude of the affective states elicited vary but can be modeled probabilistically for a given stimulus [3]. Processed sets of multimedia documents specifically prepared for controlled stimulation of emotions in laboratory settings are stored in affective multimedia databases along with their additional semantic, emotion, and context descriptors [4,5]. Based on their intended use, these documents are often referred to as stimuli, while pictures and videos are commonly referred to as visual stimuli [5].

The affective component in such datasets is typically described using the two most common models of emotion in digital systems: the dimensional and the discrete models [6]. However, there are two problems that complicate the retrieval, generation, modification, and practical implementation of affective multimedia databases: (1) the procedure for generating document emotion descriptors (i.e., ratings) requires a standardized psychological experiment conducted in a controlled environment with a statistically significant group of participants [7], and (2) the transformation of affective ratings from one model to another is not possible according to accepted theories of affect and their experimental validation [8,9].

In this paper, we investigate two types of relationships in standardized affective picture ratings: (1) between discrete emotions and (2) between combinations of single discrete emotions and affective dimensions to obtain stable clusters of emotion data points. As will be shown later, these relationships have not been thoroughly explored in general or with the specific affective multimedia datasets used here. Further, we used the Monte Carlo stabilized *k*-means method to identify clusters in discrete and dimensional spaces. This approach has not been used before. For the experiment, we used the most extensive available integrated affective picture dataset, whose structure and content are typical of other affective picture databases. Given the characteristics of the annotated images in the databases used, the conclusions drawn can be applied to other affective picture databases.

The remainder of this paper is organized as follows: Section 2 provides a detailed overview of the discrete and dimensional models of emotions in digital systems and how they are related. In particular, Sections 2.1 and 2.2 explain the dimensional and discrete emotion models, respectively, while Section 2.3 provides insight into current theories of affect concerning the relationship between the two models. Section 3 introduces the affective picture databases NAPS and NAPS BE used in the experiment to examine the correlations of picture stimuli with specific discrete emotions and clusters in the dimensional emotion space. Section 4 presents similar research investigating the relationship between emotion models in computer systems, including an overview of affective recommender systems. Section 5 describes the experiment conducted, the dataset, and the Monte-Carlo stabilized *k*-means clustering method. Section 6 reports the results obtained. Section 6.1 provides experimental results on the congruence between discrete emotion annotations in affective pictures, and Section 6.2 presents the results of picture clustering in the dimensional emotion space with respect to the dominant discrete emotions. The final section discusses the results, concludes the paper, and highlights the main interpretations with an outlook on future research.

2. Models of Emotion in Pictures

The inherent problem of emotion representation can be described in the simplest terms by the polarity and intensity of generated emotions [1,2]. Regarding emotion elicitation, some stimuli will provoke intense reactions, while others will produce no apparent feedback [10]. Moreover, dissimilar individuals or well-defined homogeneous groups will show characteristic responses depending on the stimulation semantics and context [10,11].

As mentioned earlier, in contemporary affective multimedia databases, information about emotions is described using at least one of the two models of affect: the discrete model and the dimensional model [6–8]. These two models can effectively describe emotions in digital systems but are not mutually exclusive or incompatible. While most repositories use

only one model, usually only the dimensional, some include annotations in both emotion models or offer the discrete model as an add-on option. The availability of annotations under both models is useful because it allows for a better characterization of multimedia affect content.

Affective multimedia databases have many practical applications. In addition to the study of human emotion mechanisms, the generation, and appraisal of emotions, they are employed in the study of the mutual influence of perception, memory, attention, emotion, and reasoning [12]. From a practical standpoint, they are successfully used in the detection and treatment of stress-related mental disorders such as anxiety, various phobias, and post-traumatic stress disorder (PTSD), as well as in the selection and training of personnel for occupations regularly involving high levels of stress such as in the military (for instance, [13–15]). In these applied systems, the emotion-provoking stimuli are typically delivered to participants using head-mounted displays (HMDs) or fully immersive VR [16–18]. In addition, multimedia stimuli have many different uses in different disciplines, such as cognitive science, psychology, and neuroscience, as well as in human–computer interaction and other interdisciplinary studies (for example, [19–21]).

Many affective multimedia databases currently exist. Some have general semantics, while others include content suitable for domain use, such as pictures of food, pictures in military settings, pictures with emotions rated by children only, etc. However, the Nencki Affective Picture System (NAPS) [22], the International Affective Picture System (IAPS) [23], the Geneva affective picture database (GAPED) [24], the Open Affective Standardized Image Set (OASIS) [25], and The DIsgust-RelaTed-Images (DIRTI) database [26] stand out as the most often used and the largest general picture repositories, respectively. A concise overview of these specific types of databases is provided in the related literature [4,5].

2.1. The Dimensional Emotion Model

The dimensional model of emotion is often referred to as the circumplex model of emotion, the Russell model of emotion, or the PAD (Pleasure–Arousal–Dominance) model [27–29]. It is the most commonly used emotion model for annotating multimedia stimuli. The dimensional theory of emotion assumes that a small number of dimensions can well characterize affective meaning. Therefore, dimensions are selected for their ability to statistically represent subjective emotional ratings with the smallest possible number of dimensions [23]. Russell estimated the approximate central coordinates of certain discrete emotions in the dimensional model space [27–29]. He hypothesized that these locations are not fixed but change over the life course and differ from person to person or between homogeneous groups of people based on their character traits.

The dimensional model is built around three mutually orthogonal emotion dimensions: valence (*Val*), arousal (*Ar*), and dominance (*Dom*). Positivity and negativity of a stimulus are specified by valence, while arousal describes the intensity or energy level, and dominance represents the controlling and dominant nature of the emotion. Frequently, only the first two are used because dominance is the least informative measure of the elicited affect [23]. The three dimensions are described with continuous variables *val*, *ar*, *dom*, where $val \in [1,9] \in Val$, $ar \in [1,9] \in Ar$, and $dom \in [1,9] \in Dom$, respectively. In some databases, emotion values are scaled and represented in percentages, or authors use a smaller Likert scale, e.g., in the range $[1,7]$. Thus, in the dimensional emotion model, a single emotionally annotated picture can be projected onto the two-dimensional emotion space $\Omega_{Emo} = Val \times Ar$, as exemplified in Figure 1, with each data point representing one picture from the Open Affective Standardized Image Set (OASIS) database containing 900 picture stimuli [25].
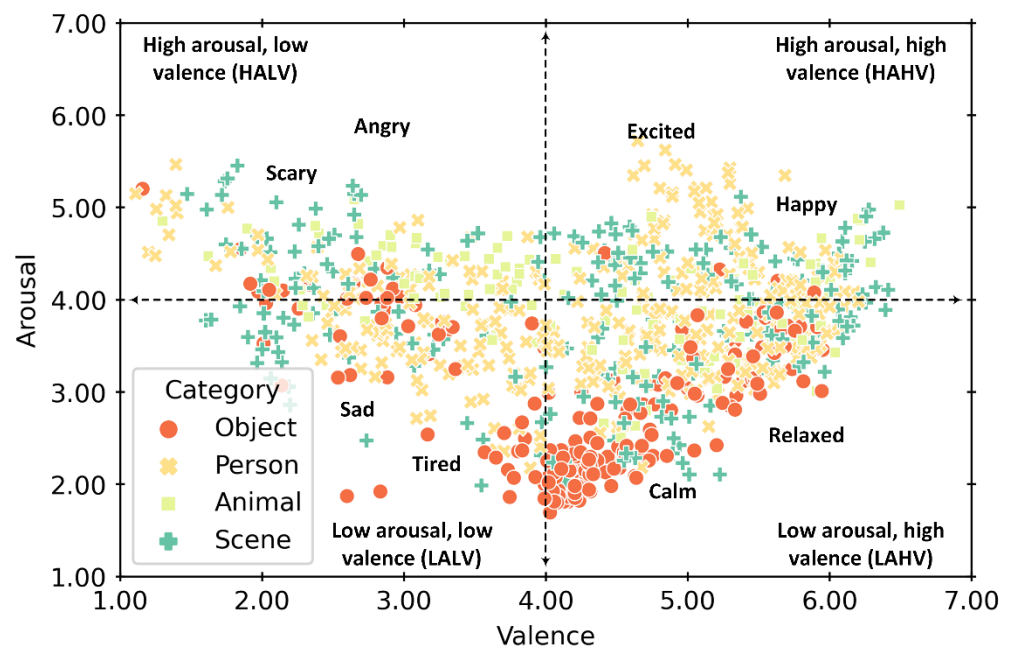
Figure 1. Picture stimuli from the OASIS dataset as an example of the dimensional emotion model (measured on a 1–7 Likert scale) with valence on the *x*-axis and arousal on the *y*-axis. The markers denote four different image categories: object, person, animal, and scene. Approximate positions of discrete emotions are indicated. For more information on the dimensional model, see [28–30]. Reproduced with permission from Kurdi, B.; Lozano, S.; Banaji, M.R., Introducing the open affective standardized image set (OASIS); published by Springer, 2017.

The distance between two emotions expressed in the dimensional space is Euclidian. Therefore, the similarity between two dimensional emotions a^{dim} and b^{dim} with each emotion vector containing components of valence, arousal, and dominance may be expressed as:

$$sim\left(a^{dim}, b^{dim}\right) = 1 - dist\left(a^{dim}, b^{dim}\right)$$
$$dist\left(a^{dim}, b^{dim}\right) = \sqrt{\left(a_{val}^{dim} - b_{val}^{dim}\right)^2 + \left(a_{ar}^{dim} - b_{ar}^{dim}\right)^2 + \left(a_{dom}^{dim} - b_{dom}^{dim}\right)^2} \quad (1)$$

2.2. The Discrete Emotion Model

The discrete or categorical model classifies emotions into specific labels, referred to as basic emotions, discrete emotions, emotion norms, or norms for short [31]. Categorical emotion theories claim that dimensional models do not accurately reflect the neural systems underlying emotional responses [8]. Proponents of these theories propose instead that there is a universal set of emotions [8]. However, the exact number of basic emotions is a point of dispute among experts. Contemporary psychological theories propose discrete emotion models with up to 24 different norms [6]. Nevertheless, most researchers agree that six primary emotions—happiness, surprise, fear, sadness, anger, and disgust—are universal across cultures and have an evolutionary and biological basis [32]. According to the universality hypothesis, these six basic internal human emotions are expressed with the same facial movements in all cultures, which supports a universal recognition of emotions in facial expressions. Moreover, some experimental evidence exists for mixed emotion models, at least during some life stages [33]. However, in all models, the simultaneous presence of more than one discrete emotion is allowed but with different intensities. The emotion with the highest intensity is called the dominant emotion. A clear example of how a discrete model is used in computer systems is the NimStim stimuli set of facial expressions [34] as illustrated in Figure 2. The NimStim picture dataset was created to provide facial expressions that untrained individuals would recognize as typical of

research participants. This set is comprehensive, multiracial, and available online to the scientific community together with the accompanying pictures' normative ratings [34]. In general, datasets similar to NimStim use the discrete model. They are important because behavior, cognition, and emotion experiments often rely on highly standardized picture datasets to understand how people respond to different faces that convey visually specific, discrete emotions of varying intensity. These datasets have played an essential role in the development of face science [35].

Figure 2. Examples of 6 discrete emotions represented in the NimStim picture database [34]. From top to bottom, and left to right: happiness, sadness, surprise, anger, fear, disgust. Adapted from [34].

Individual norms can be mathematically described as mutually orthogonal unit vectors that collaboratively define an *n*-dimensional emotion space. Intensity of each vector is a dimensionless value normalized in the range [0.0, 100.0]. Therefore, a discrete emotion a^{dis} in a 6-component emotion space defined with norms of happiness, sadness, surprise, anger, fear, disgust is a vector: $a^{dis} = \left[a^{dis}_{happiness}, a^{dis}_{sadness}, a^{dis}_{surprise}, a^{dis}_{anger}, a^{dis}_{fear}, a^{dis}_{disgust} \right]^{T}$. Consequently, given two discrete emotions a^{dis} and b^{dis} where each contains *n* components, the distance $dist\left(a^{dis}, b^{dis} \right)$ between them can be written as:

$$dist\left(a^{dis}, b^{dis} \right) = 1 - sim\left(a^{dis}, b^{dis} \right)$$
$$sim\left(a^{dis}, b^{dis} \right) = \cos(\Theta) = \frac{a^{dis} \cdot b^{dis}}{\|a^{dis}\| \, \|b^{dis}\|} = \frac{\sum_{i=1}^{n} a^{dis}_i b^{dis}_i}{\sqrt{\sum_{i=1}^{n} a^{dis}_i} \sqrt{\sum_{i=1}^{n} b^{dis}_i}} \tag{2}$$

The distance $dist\left(a^{dis}, b^{dis} \right)$ is a numerical value in the range [0.0, 1.0].

2.3. The Relationship between Dimensional and Discrete Models

The traditional view of the relationship between affective dimensions and emotional category ratings is based on previous experiments characterizing affective stimuli: written words [36], facial expressions of emotion [37], and affective sounds [38]. These experiments show differences in predictions based on categories depending on the predicted dimension

and whether the pictures were positive or negative. The regressions using the dimensional ratings to predict the emotional category ratings are not homogeneous concerning the ability of the categorical ratings to predict the dimensional ratings. In other words, according to currently accepted theories and their experimental validations, emotion categories cannot be extrapolated from affective dimensions; conversely, dimensional information cannot be inferred or extrapolated from emotion categories. The heterogeneous relationships between each emotion category and the different affective dimensions of the visual stimuli confirm the importance of using categorical data independently and as a complement to dimensional data [38]. From a practical perspective, by using both dimensional and discrete emotion classifications, the researcher could design a more ecologically valid paradigm [38].

3. The NAPS and NAPS BE Affective Picture Databases

Affective multimedia databases are standardized digital repositories that store auditory, linguistic, and visual materials for emotion research [5]. Currently, there are many databases indexing affective information in multimedia [4,5], but to our knowledge, the Nencki Affective Picture System (NAPS) is currently the largest database of visual stimuli with general semantic content and with a comprehensive set of accompanying normative ratings and physical parameters of pictures (e.g., luminance, contrast) [22,39].

The NAPS picture system is the result of a multi-year research effort conducted by the Polish Nencki Institute of Experimental Biology [22]. This database was constructed in response to limitations of existing databases, such as a limited number of stimuli in specific categories or poor picture quality of visual stimuli. The NAPS can be used in various affective research areas. The original database comprises 1356 realistic, high-quality photographs divided into five disjoint categories: people, faces, animals, objects, and landscapes. The photos were chosen to elicit a specific emotional response in the general population, i.e., not all photos are content neutral. One of the main features of the NAPS picture system as a whole, i.e., with its extensions, compared to other emotion elicitation repositories, is that it combines a relatively large number of pictures with normative ratings elicited per both dimensional and categorical (discrete) emotion theories [22], in addition to having additional multi-word semantic descriptions organized in a variety of different topics.

More recently, three extensions of NAPS have been developed. The first is NAPS Basic Emotions (NAPS BE), which includes normative ratings based on the discrete model of emotions and additional dimensional ratings for a subset of 510 images from the original NAPS [9]. These additional ratings were obtained using a different psychometric methodology than that NAPS [9]. It is important to emphasize that NAPS BE does not contain additional images but only discrete and additional dimensional emotion annotations to the original 1356 pictures in NAPS. Examples of 6 pictures (Animals_073_h, Animals_177_h, Faces_116_h, Faces_192_h, Landscapes_025_h, Landscapes_121_h) with dimensional emotion ratings from NAPS are shown in Figure 3. Mean values and standard deviations of discrete emotions ($happiness\_M$, $sadness\_M$, $fear\_M$, $surprise\_M$, $anger\_M$, $disgust\_M$, $happiness\_SD$, $sadness\_SD$, $fear\_SD$, $surprise\_SD$, $anger\_SD$, $disgust\_SD$) from NAPS BE for these 6 pictures are shown in the chart in Figure 4. For further illustration of the relationship between dimensional and discrete models in affective images, see [40].

The dimensional picture ratings of NAPS were collected using Likert-scale questionnaires, while participants in the creation of NAPS BE provided their subjective ratings using a standardized Self-Assessment Manikin (SAM) graphic scale [41,42]. The SAM is a non-verbal pictorial assessment technique that directly measures the pleasure, arousal, and dominance associated with a person's affective reaction to a wide variety of stimuli [41,42]. It has been shown that SAM provides consistent and culturally independent cues to emotional responses [41,42]. The SAM can be used in a variety of settings with different populations, adults, and children. Because of its brevity, it can be used to rapidly assess emotional responses to a wide range of emotion elicitation methods [41,42]. To acquire NAPS and NAPS BE picture ratings, several hundred participants from several countries representing the general population provided input in a controlled environment.
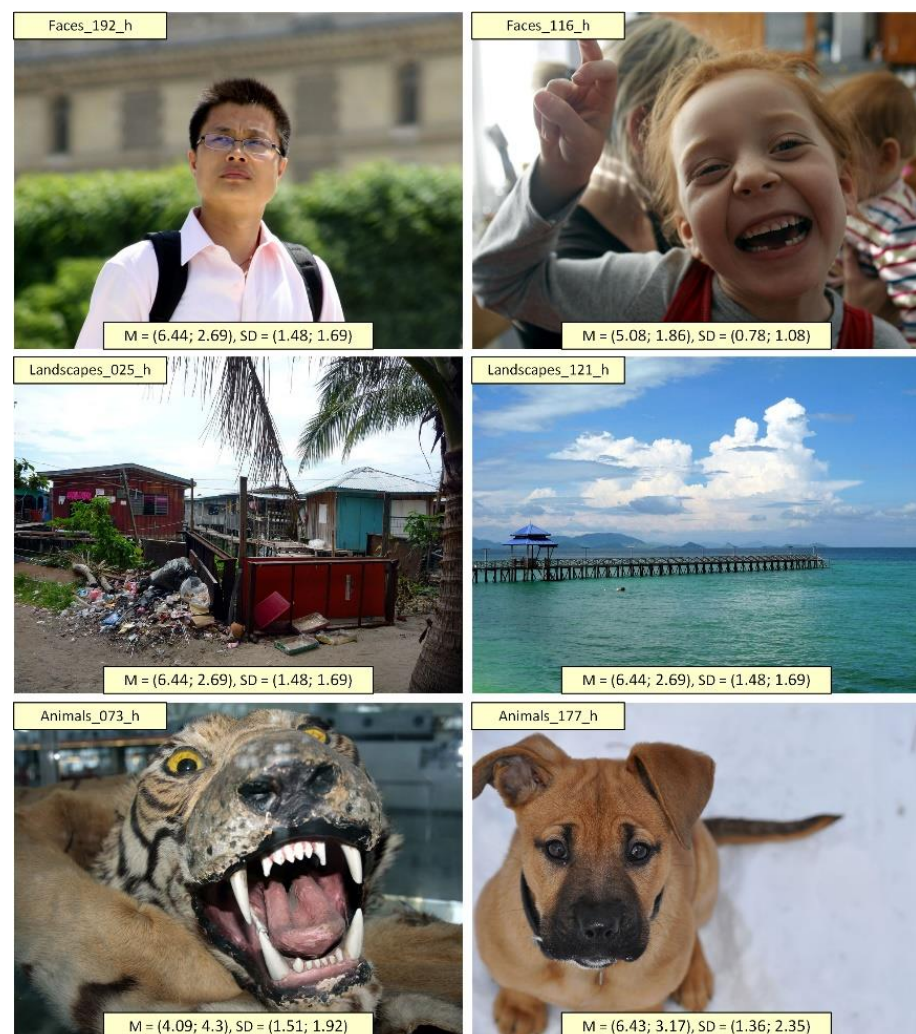
Figure 3. Examples of 6 pictures from the NAPS repository used in the experiment annotated with their dimensional emotion values: mean (*val_SAM_M*, *aro_SAM_M*) and standard deviation (*val_SAM_SD*, *aro_SAM_SD*). From top to bottom row, the pictures belong to the subgroups faces, landscape, and animals, respectively. Reproduced with permission from [22].
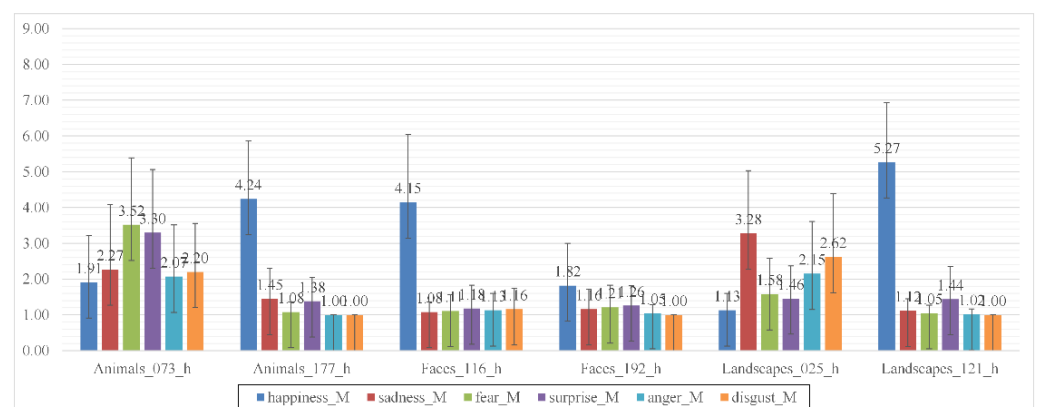

Figure 4. Mean and standard deviations values (*happiness_M*, *sadness_M*, *fear_M*, *surprise_M*, *anger_M*, *disgust_M*, *happiness_SD*, *sadness_SD*, *fear_SD*, *surprise_SD*, *anger_SD*, *disgust_SD*) from NAPS BE database representing discrete emotions of 6 pictures used in the experiment. The pictures are illustrated in Figure 3.

In later efforts to expand NAPS, an erotic subgroup was introduced for the Nencki Affective Picture System (NAPS ERO) with an additional 200 stimulus pictures that were provided with self-reported ratings of emotional valence and arousal by homosexual and heterosexual men and women ($N = 80$, divided into four equally sized subgroups) [43]. Finally, Children-Rated Subset is the most recent extension of NAPS and includes 1128 images from the original database that were rated as appropriate for children based on various criteria and expert judgment. In the Children-Rated Subset, affective ratings were collected from a sample of $N = 266$ children aged 8–12 years [44].

In our experiment, we used the NAPS and its first extension NAPS BE. Both databases have a typical architecture and file structure and use a data model standard of most affective picture databases [5]. Their combined dataset is sufficiently large to make inferences about the relationship between affective variables. Moreover, they were developed relatively recently. As such, they are the best common representatives of affective image sets using the dimensional and basic models of emotion. Therefore, we believe that these two data sets were the optimal choice for the presented study.

4. Related Work

The relationship between separate emotional states and their influence on different mental disorders such as depression has been the topic of much research covering various modalities of emotion representation, expression, and recognition such as speech and facial expressions. However, the investigation of transformations between discrete and dimensional emotion models in affective multimedia databases using machine learning methods is more sparse.

A method based on cluster analysis was proposed to improve the selection of IAPS stimuli [45]. Three clustering procedures—k-means, hierarchical, and model-based clustering—were used to produce a set of coherent clusters. This study aimed to identify the most likely group structures based on valence, arousal, and dominance norms criteria. Only stimuli similarity in the dimensional model were considered [45]. The MuSe-Toolbox (The Multimodal Sentiment Analysis Continuous Annotation Fusion and Discrete Class Transformation Toolbox) is a novel Python-based open-source toolkit for emotion recognition [46]. The toolkit extracts complex time-series features from audiovisual and psychological signals and processes them with gold standard methods in machine learning, such as regression and classification. With the toolkit, it is possible to automatically cluster continuous signals and transform them in discrete emotions [46]. In [47], clustering algorithms were used to classify affective pictures representing objects based on valence and arousal dimensions to examine participants' depression traits. It was shown that the depression trait does not significantly affect the accuracy or time-order of emotional classification. In another study, global image properties such as local brightness contrast, color or the spatial frequency profile were used to predict ratings of affective pictures [48]. This study investigated whether image properties that reflect global image structure and image composition affect the rating of visual stimuli from IAPS, NAPS, OASIS, GAPED, and DIRTI affective picture databases. An SVM-RBF classifier was used to predict high and low ratings for valence and arousal, respectively, and achieved a classification accuracy of 58–76%. Additionally, a multiple linear regression analysis showed that the individual image properties account for between 6 and 20% of the variance in the subjective ratings for valence and arousal. In an empirical research $N = 70$ healthy subjects mapped discrete emotion labels to PAD space according to their personal understanding using a simple software tool [49]. There was a high inter-subject consistency regarding the positioning of discrete emotions for the dimension of valence. However, arousal and dominance ratings showed considerably greater variance. It was concluded that global and reliable mappings of discrete emotions into the dimensional emotion space can best be provided for the valence dimension. The CAKE model was developed to comprehensively analyze how Deep Neural Networks (DNN) can be used to represent emotional states [50]. For this purpose, researchers studied how many dimensions are sufficient to accurately represent

an emotion resulting from a facial expression. It was concluded that three dimensions are a good trade-off between accuracy and compactness, agreeing with the dimensional emotion model of valence–arousal–dominance.

Finally, in a report on the application of the dimensional emotion model in automatic emotional speech recognition, scholars used the perceptron rule in combination with acoustic features from speech to classify the utterance in valence–arousal emotion space [51]. They used two corpora of acted emotional speech: the Berlin Emotional Speech Database (in German) and the Corpus of Emotional and Attitude Expressive Speech (in Serbian language). Their experimental results showed that the discrimination of emotional speech along the arousal dimension is better than the discrimination along the valence dimension for both corpora.

Compared to the previously published work, our research uses a novel experimental setup to obtain transformations between discrete and dimensional emotion spaces in affective picture databases. Additionally, we obtained strong correlations in both emotion spaces.

Affective Recommender Systems

In general, recommender systems (RSs) are software tools and techniques that provide suggestions for items that are most likely of interest to a particular user [52]. There are many references in published literature to recommender systems being used in human–computer interaction (HCI) using information about emotions to leverage multimedia content (e.g., music, video clips, text, and images) and help users to discover interesting new items based on their personal preference features [53]. Depending on the application in question, the features extracted from media items may be semantically oriented into prediction of particular target classes in the expectation that this will eventually have a positive impact on the final quality of recommendation. For instance, features for an automatic playlist generation system extracted from music items may be oriented into prediction of some predefined emotion or genre classes, human-generated tags (auto-tagger), or some latent factor features of collaborate models.

Affective recommender systems (ARSs) are a type of RSs that make recommendations based on estimation of emotion using multimodal data fusion of various physiological signals, facial expressions, body posture, gestures, eye movements, voice, and speech [54]. The emergence of sensors and wearable devices as mechanisms for capturing physiological data from people in their daily lives has enabled research on detecting emotional patterns to improve user experience in different contexts. For example, in [55], the authors apply the deep learning approach to a dataset of physiological signals—electrocardiogram (ECG) and galvanic skin response (GSR)—to detect users' emotional states. In this study, relevant features are extracted from the physiological signals using a Deep Convolutional Neural Network (DCNN) to perform emotion recognition. The estimated valence and arousal values can be used to classify the affective state and as input to personal recommendation systems [55]. Further, it has been realized that emotions are important information to consider when suggesting instructional content to students in a learning environment, since it might change their emotional state [56]. To examine the current state of the influence of emotions in the field of education, particularly in content recommender systems, a systematic literature review of affective recommender systems in learning settings was conducted [56]. Finally, research should be noted concerning a video streaming service that uses the viewer's emotional reactions as the basis for recommending new content [57]. The researchers conducted a study where subjects' facial expressions and skin-estimated pulse were monitored while watching videos. Results of this study showed 70% accuracy in estimation of dominant emotions. However, no correlation was found between the number of emotional reactions participants have and how they rate the videos they watch. The pulse estimation was reliable to measure important changes in pulse [57].

The approach used in our work may be integrated into an affective recommender system. However, unlike the research on affective recommender systems mentioned above,

our work focuses specifically on developing recommender systems for the construction of multimedia stimuli databases. Although several subtypes of RSs exist [52], such as collaborative filtering (CF), context-based, content-based, and session-based RSs, our approach would likely use only the content-based type of RS. Here, an image stimulus would be recommended for display to a user based on the similarity of the image in a constructed cluster (using the *k*-means algorithm) to an image previously shown to the user. The clusters are determined later in the paper and depend on the emotion pairs evoked by the image. To address the cold start problem, the first image recommended to a user would require initial input from the user regarding the emotional states the image should evoke (e.g., the user wants an image that makes them feel surprised and fearful), which is a common approach that falls under the umbrella of active learning [58]. An activity schema of the proposed recommender system for iterative construction of affective picture databases using relationships between discrete and dimensional emotion models is shown in Figure 5.
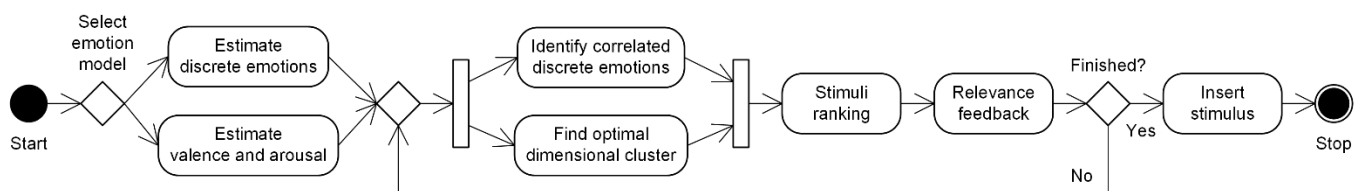

Figure 5. UML activity diagram of the proposed recommender system for iterative construction of affective picture databases using relationships between discrete and dimensional emotion models.

In the proposed recommender system, after the initial estimation of discrete or dimensional emotions is provided, correlated discrete emotions and optimal clusters of dimensional values are identified based on the results of our work. Then, the user provides relevance feedback through the system's graphical user interface (GUI) and improves stimuli emotion classification. If the result is not satisfactory, the whole loop is repeated. Finally, the stimulus is inserted in the affective picture database when the user has decided that the picture stimulus has acceptable discrete and dimensional emotion annotations.

5. Experimental Setup

In the experiment, we have comprehensively evaluated relationships between picture stimuli in the NAPS and NAPS BE datasets based on the discrete and dimensional emotion models. The experiment is divided into two main segments. In the first segment, we looked at discrete emotions in NAPS BE and explored correlations between them. We correlated these discrete emotion values with the supplementary dimensional ratings in the NAPS BE. In the second segment of the experiment, using the NAPS dataset, we performed clustering for the most correlated tuples of discrete and discrete-dimensional emotions from the first part.

A software tool was developed for the experiment in the Python programming language. The tool uses the scikit-learn, numpy, scipy, and matplotlib libraries for the implementation of machine learning algorithms, data exploration, and visualization, respectively. The software tool and experiment results are freely available for scientific and non-commercial purposes at the following URL: https://github.com/kburnik/naps-clustering (accessed on 18 July 2022). For all inquiries, please contact the third author. The archive does not contain NAPS and NAPS BE. To request these datasets for non-profit academic research purposes, please contact Nencki Institute of Experimental Biology, Laboratory of Brain Imaging (LOBI) at https://lobi.nencki.gov.pl/research/8/ (accessed on 18 July 202).

5.1. Dataset

As described in Section 3, the NAPS BE is a data description extension to the NAPS and consists of 510 normative picture annotations containing attributes representing 6 underlying discrete emotions or norms: happiness, fear, sadness, surprise, anger, and

disgust [9,22]. In this respect, additionally to the NAPS dataset, each picture stimulus stim is described with 6 mean values of discrete emotions (*happiness_M*, *sadness_M*, *fear_M*, *surprise_M*, *anger_M*, *disgust_M*) and the corresponding 6 standard deviation values (*happiness_SD*, *sadness_SD*, *fear_SD*, *surprise_SD*, *anger_SD*, *disgust_SD*). Further, the NAPS BE dataset provides 4 values representing dimensional emotion variables valence and arousal (*val_SAM_M*, *aro_SAM_M*, *val_SAM_SD*, *aro_SAM_SD*) supplementary to the values of these variables in the NAPS dataset. The latter values for valence and arousal are different for the same picture in the two datasets because they were obtained using different experimentation methodologies, as explained in the previous section. Thus, the study dataset has a total of 16 unique attributes.

5.2. Monte-Carlo Stabilized k-Means Clustering

The *k*-means centroid algorithm is well-known and often used by researchers in different domains. The most important drawback of the algorithm is its dependence on the initial conditions, which produces statistical distribution instability. However, the nondeterministic behavior of the *k*-means algorithm was solved by Monte-Carlo simulation [59] with the selection of the highest frequencies of affiliation in the distributions so that the results could be reproduced without introducing arbitrary initial conditions. The silhouette method was used to test the robustness of the partitions, i.e., the separation distance between the resulting clusters in the dimensional emotion space. This method attempts to estimate for each data point how strongly it belongs to the assigned cluster and, at the same time, how weakly it belongs to other clusters [60]. The corresponding silhouette plot shows how close each point in a cluster is to points in neighboring clusters, providing a way to visually evaluate parameters such as the number of clusters. Our previous research shows that the optimum distribution of pictures stimuli in the NAPS dimensional space is achieved for $k = 4$–7 clusters [39]. This finding was used in the second segment of the experiment to limit the search space.

6. Results

The results of the first segment of the experiment investigating the congruency in stimuli emotion norms and dimensions are described in Section 5.1, and the second part of the experiment, which explores the clustering of picture stimuli in the dimensional space based on correlations between discrete and dimensional ratings, is brought forward in Section 5.2.

6.1. Data Congruency in Emotion Norms

As the first step in the experiment, distinctive pairs of discrete emotions from the NAPS BE dataset are considered to establish a correlation between them. In this regard, Pearson (r) and Spearman's rank correlation (r_s) coefficients were obtained for 15 distinct pairs of 6 emotion norms. The results are shown in Figure 6.

The results show that a moderate negative correlation can be observed for norm pairs {happiness, sadness}, {happiness, anger}, and {happiness, disgust} with $r = -0.67$, -0.62, -0.62, respectively. Additionally, there is a strong positive correlation of {sadness, anger} 0.81, {fear, surprise} 0.76, and of {surprise, disgust} 0.71, as detailed in Figure 7. The highest Spearman's rank coefficient is observed for three tuples {sadness, anger}, {happiness, sadness} and {happiness, anger} with $r_s = 0.85$, -0.81, and -0.81, respectively. This indicates strong monotonic relationships where stimuli with a high value of sadness also have high value of anger, and oppositely, when happiness is high, sadness and anger are low. Interestingly, the relationship in tuples {fear, anger} and {anger, disgust} is more continuously growing than in the other pairs. Some negative polarity emotions are mutually less co-linear in absolute terms than they are with the positive emotion of happiness. Likewise, although negative polarity emotions are not all strongly co-linear, sadness and anger tend to increase together the most of all discrete emotion pairs.
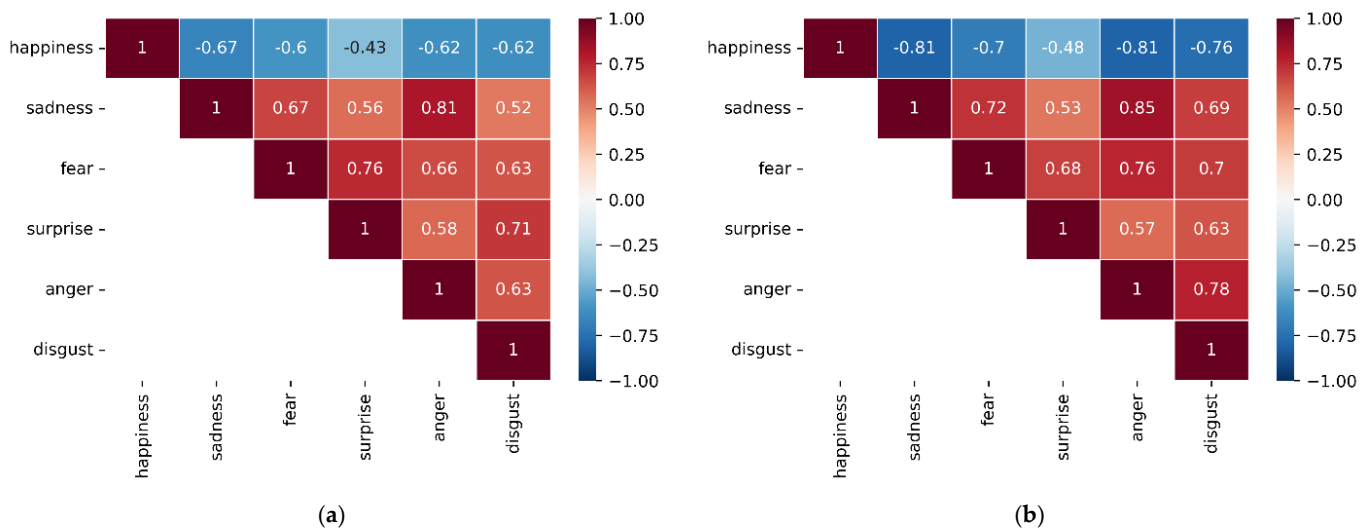
(a)

(b)

Figure 6. Correlations between 6 emotion norms (happiness, fear, sadness, surprise, anger, disgust) in the NAPS BE dataset: (**a**) Pearson correlation coefficient matrix (r); (**b**) Spearman's rank matrix (r_s).
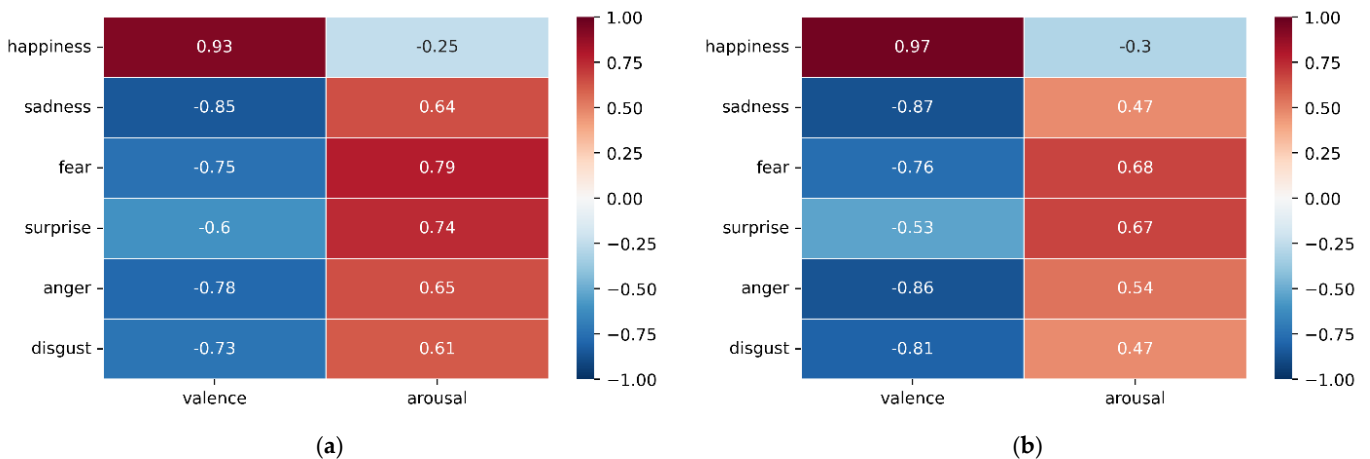

(a)

(b)

Figure 7. Correlations among 6 emotion norms (happiness, fear, sadness, surprise, anger, disgust) and 2 emotion dimensions (valence, arousal) in the NAPS BE dataset: (**a**) Pearson correlation coefficient matrix (r); (**b**) Spearman's rank matrix (r_s).

Having the correlations between discrete emotion pairs, the same analysis is performed for all 12 possible combinations of discrete emotions and emotion dimensions. As can be seen in Figure 7, there is a very strong positive correlation between happiness and valence ($r = 0.93$). Strong positive correlations also exist for fear and arousal ($r = 0.79$), and surprise and arousal ($r = 0.74$). On the other hand, strong negative correlations are found between valence and four negative polarity norms: sadness, fear, anger, and disgust, the highest being for sadness ($r = -0.85$) and anger ($r = -0.78$). The findings confirm previously reported results about the relationship between different discrete emotions in affective processing [61]. These observations highlight interesting pairs of basic emotions for further in-depth analysis regarding clustering in dimensional space.

6.2. Picture Clustering Based on Dominant Discrete Emotions

In the second segment of the experiment, we investigated how basic emotion ratings are associated with well-defined regions in dimensional emotion space. As previously explained, we identified the optimal number of homogeneous clusters of dimensional emotions in the NAPS dataset [39]. Using the Monte-Carlo simulation stabilized k-means algorithm with $p = 2000$ iterations in range $k = 2$–7, we found that for $k = 4$ the stability error

is only 0.29% of the total number of the NAPS data points. These findings have provided us with the examination of their relationship with basic emotions.

As the NAPS BE dataset is more complex than the NAPS in terms of the number of data dimensions, to visualize the cluster analysis results, only statistically significant correlations of discrete and dimensional ratings were considered. Silhouette coefficients graphs for the corresponding pairs of discrete and homogeneous clusters in dimensional emotion space are shown in Figure 8. It can be noticed that the silhouette method shows that the evaluation is consistently low when there are more than 8 distributions. This shows that the range $k = 2$–7 is sufficient for mapping NAPS BE data points to the clusters in the NAPS database.
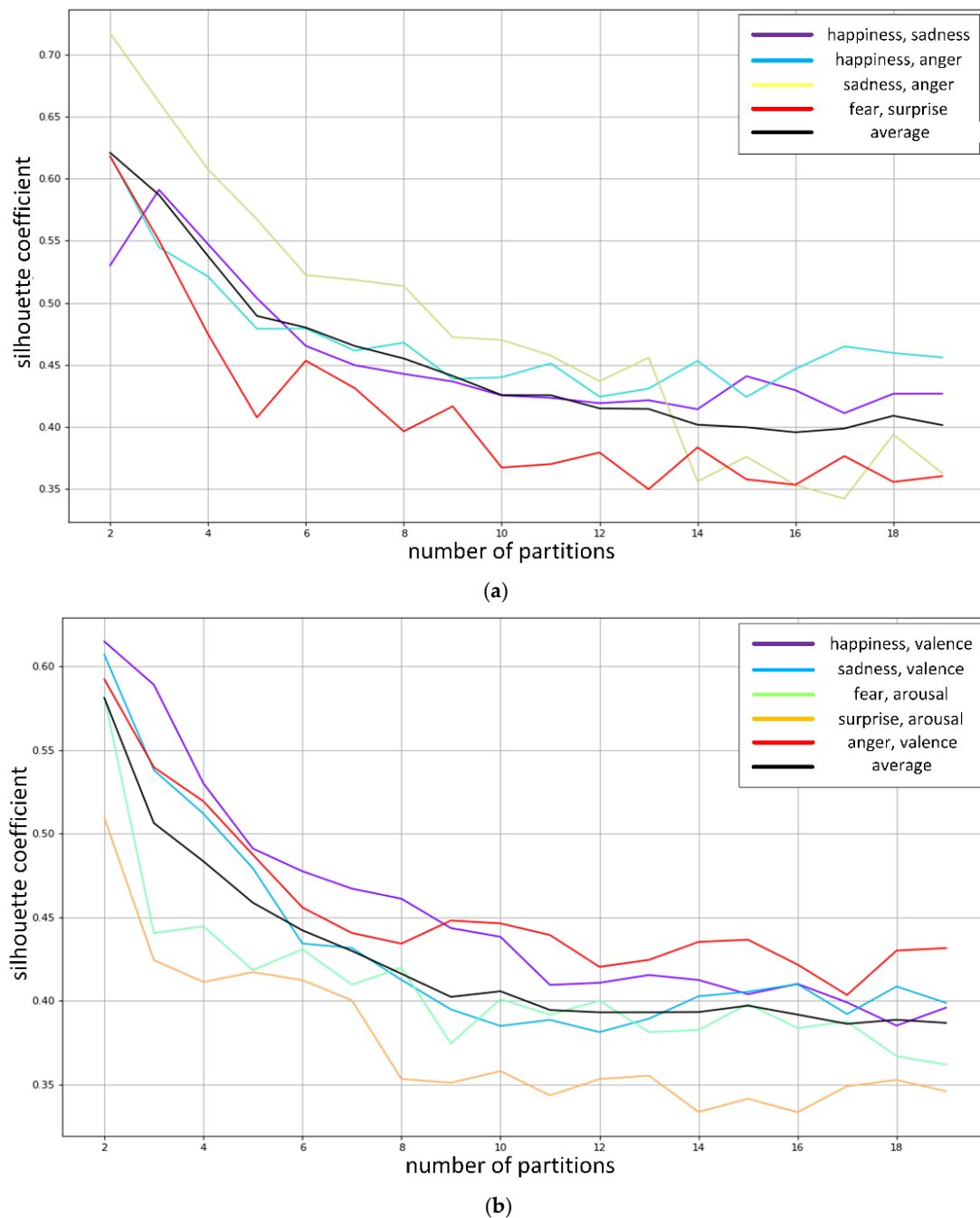

(**a**)

(**b**)

Figure 8. Silhouette plots in interval $k = 2$–19 for: (**a**) pairs of emotion norms {happiness, sadness}, {happiness, anger}, {sadness, anger}, {fear, surprise}; (**b**) tuples of emotion norms and emotion dimensions {happiness, valence}, {sadness, valence}, {fear, arousal}, {surprise, arousal}, {anger, valence}.

Additional analyzes showed that the decrease in the average value of the silhouette coefficient for $k > 4$ was less significant than the increase in the number of clusters. Therefore, for the final results of stable partitioning of isolated correlation examples from the NAPS BE database into the NAPS database, the division of dimension scores into $k = 4$ clusters were chosen. The additional fact supports this choice that NAPS BE is almost three times smaller than the NAPS database, which is reflected in the lower point density. Furthermore, the choice of $k = 4$ is supported by a selection of the smallest number of homogeneous clusters, i.e., those with the lowest scatter of all data points in the dataset, as reported in [39].

The scatter plots in Figure 9 show k-means clustering results in dimensional space with $p = 2000$ iterations using the approach stabilized by Monte-Carlo simulation [39]. The results were obtained for $k = 4$ individual clusters and matched with the following pairs of discrete and dimensional emotion data dimensions: {happiness, sadness}, {fear, surprise}, {happiness, valence}, and {surprise, arousal}. These emotion tuples provide the most robust transformations of values between the two emotion models. The Monte-Carlo stabilized k-means method proved to be highly reliable. The stability error was measured to be only 0.29% of the total number of the NAPS database data points for the chosen number of data clusters $k = 4$ and $p = 2000$ iterations.
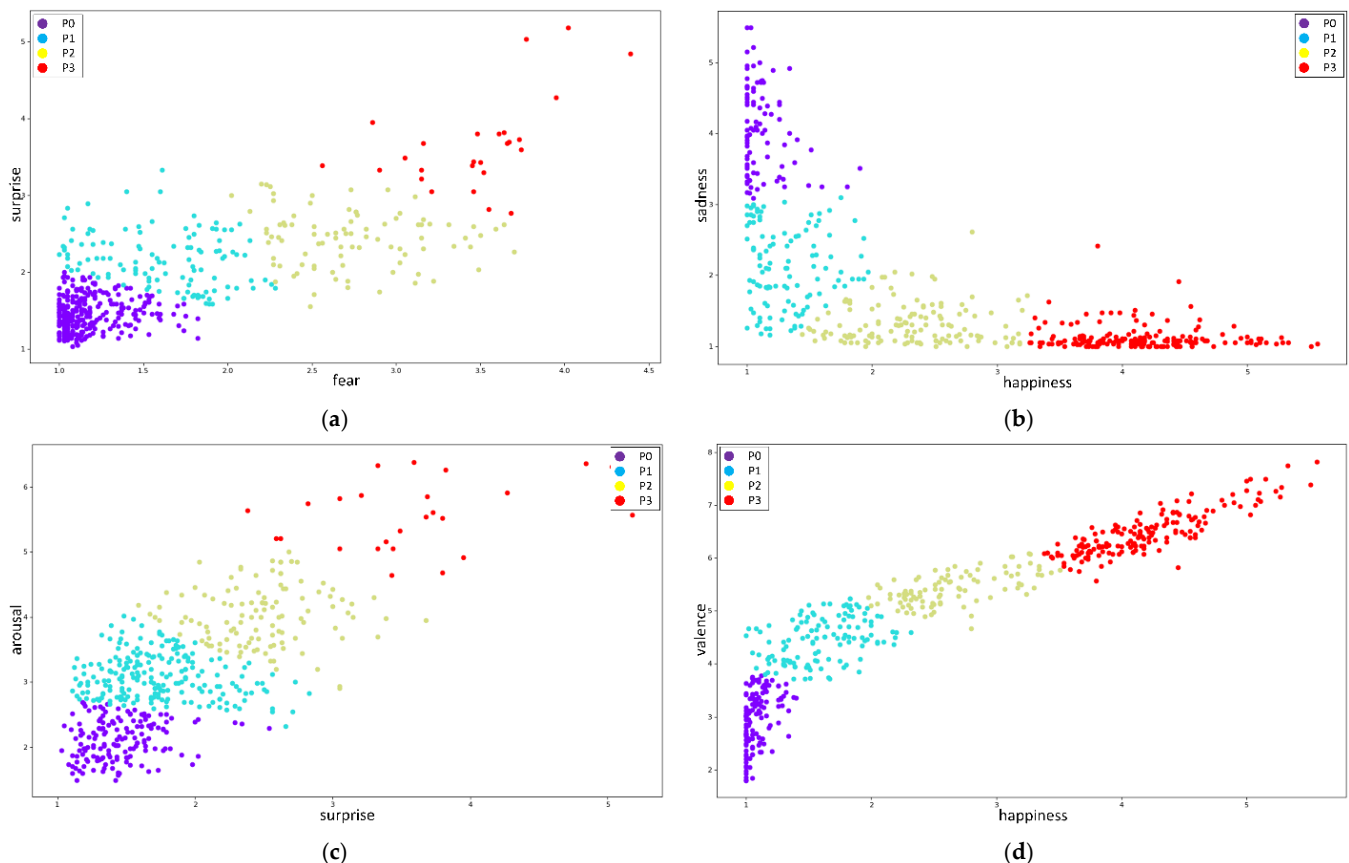

(a)　　　　　　　　　　　　　　　　　　　　　　　　(b)

(c)　　　　　　　　　　　　　　　　　　　　　　　　(d)

Figure 9. After $p = 2000$ iterations of Monte-Carlo stabilized k-means algorithm ($k = 4$), cluster distribution of transformations between discrete and dimensional models in the NAPS BE dataset was achieved for the following specific emotion tuples: (**a**) {fear, surprise}; (**b**) {happiness, sadness}; (**c**) {surprise, arousal}; (**d**) {happiness, valence}.

The results are consistent with the expected relationship between dimensional emotions and semantic picture annotations in the NAPS database. In previously reported research, the optimal number of centroids in the dimensional space related to the description of stimuli was estimated using the minimum cumulative error rule [62]. Although the exact value of clusters was not determined, it was found that the optimal number is definitely in the open interval of $k = 2$–16.

7. Conclusions

This study aimed to investigate the relationship between the two most common emotion models used in affective picture databases: the dimensional and discrete models. We did not attempt to create a new larger emotion model as a union of the two existing models but rather to explore if it is possible to identify statistically relevant dependencies between individual discrete emotions and discrete emotions and emotion dimensions.

The experiment determined a set of statistical dependencies between discrete and dimensional emotion variables using the NAPS and NAPS BE databases, containing 510 and 1356 emotionally annotated pictures, respectively. In the presented research, distance metrics in discrete and dimensional emotion space have been described.

The correlated pairs of discrete emotions in picture stimuli have been reported with the highest positive correlation between tuples of discrete emotions {anger, sadness}, {fear, surprise}, and {disgust, surprise}, and negative correlation values with {happiness, sadness} and {happiness, anger}. Significant positive correlation was observed between discrete and dimensional emotions tuples {happiness, valence}, {arousal, fear}, and {arousal, surprise}. Media with these predominant discrete and dimensional emotions are mutually complementary in eliciting specific affective states. Using these findings, researchers who employ affective multimedia databases, such as psychologists, psychiatrists, and neurologists, may select the best available appropriate multimedia stimuli if the optimal picture stimulations are not accessible.

Nevertheless, the selection of stimuli based only on emotional content is insufficient in provoking desired emotional reactions. Therefore, the semantic meaning of multimedia must be selected with the greatest care so that the stimuli are effective and have the desired personal importance and ego relevance to the participants of an emotion elicitation experiment [63].

The NAPS and NAPS BE selected for the study are the largest general affective picture databases with dimensional and discrete annotations available. However, a validation experiment using more recent databases in the future should also be considered to confirm the reported findings, such as the OASIS [25].

There are many possible practical uses of the presented research. Ultimately, we would like to implement the attained findings in an intelligent system that can reason about the emotional content of a multimedia document. Transforming information about emotions between different models would help to discover hidden knowledge in multimedia documents. This would facilitate the generation of additional affective picture annotations and the construction of a more detailed and comprehensive description of database content. Such an intelligent system could potentially have many useful applications, such as supported construction of affective multimedia databases, video recommendation, or emotion estimation. Finally, the presented research has not investigated complex relationships between emotional norms and multimedia semantics.

The supported construction of affective multimedia databases is much needed and will lead to better utilization in practice. Because current construction processes require the evaluation of many subjective self-reports collected in stimuli annotation experiments, the workload and resources needed to perform these tasks hinder the employment of affective multimedia databases. Therefore, automation of the stimuli annotation process is necessary to facilitate the development of affective multimedia databases and propagate their usage. The congruency of semantics and emotion allows us to obtain a priori knowledge about stimulus semantics to derive its emotion ratings using semantic and emotion annotations stored in the affective multimedia databases [63]. In such a construction process, the congruency may be used to retrieve a set of the most likely emotion annotations for a new multimedia stimulus. The results would be assessed, sorted by estimation probability, and displayed to the user. Interactive user feedback can be used to optimize or validate classification through a sequence of questions and answers between the stimulus annotation system and the human annotation expert [63]. This exciting direction of research will be our future focus as well.

References

1. Scherer, K.R. What are emotions? And how can they be measured? *Soc. Sci. Inf.* **2005**, *44*, 695–729. [CrossRef]
2. Lewis, M.; Haviland-Jones, J.M.; Barrett, L.F. *Handbook of Emotions*; Guilford Press: New York, NY, USA, 2010.
3. Picard, R.W. Affective computing: Challenges. *Int. J. Hum.-Comput. Stud.* **2003**, *59*, 55–64. [CrossRef]
4. Wang, Y.; Song, W.; Tao, W.; Liotta, A.; Yang, D.; Li, X.; Gao, S.; Sun, Y.; Ge, W.; Zhang, W.; et al. A systematic review on affective computing: Emotion models, databases, and recent advances. *Inf. Fusion* **2022**, *83–84*, 19–52. [CrossRef]
5. Horvat, M. A brief overview of affective multimedia databases. In Proceedings of the Central European Conference on Information and Intelligent Systems, Varaždin, Croatia, 27–29 September 2017; Faculty of Organization and Informatics: Varaždin, Croatia, 2017; pp. 3–9.
6. Horvat, M.; Stojanović, A.; Kovačević, Ž. An overview of common emotion models in computer systems. In Proceedings of the 45th Jubilee International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO 2022), Opatija, Croatia, 23–27 May 2022; pp. 1152–1157.
7. Colden, A.; Bruder, M.; Manstead, A.S. Human content in affect-inducing stimuli: A secondary analysis of the international affective picture system. *Motiv. Emot.* **2008**, *32*, 260–269. [CrossRef]
8. Peter, C.; Herbon, A. Emotion representation and physiology assignments in digital systems. *Interact. Comput.* **2006**, *18*, 139–170. [CrossRef]
9. Riegel, M.; Żurawski, Ł.; Wierzba, M.; Moslehi, A.; Klocek, Ł.; Horvat, M.; Grabowska, A.; Michałowski, J.; Marchewka, A. Characterization of the Nencki Affective Picture System by discrete emotional categories (NAPS BE). *Behav. Res. Methods* **2016**, *48*, 600–612. [CrossRef] [PubMed]
10. Coan, J.A.; Allen, J.J. *Handbook of Emotion Elicitation and Assessment*; Oxford University Press: Oxford, UK, 2007.
11. Popović, S.; Horvat, M.; Kukolja, D.; Dropuljić, B.; Ćosić, K. Stress inoculation training supported by physiology-driven adaptive virtual reality stimulation. *Annu. Rev. Cyberther. Telemed.* **2009**, *144*, 50–54.
12. Brosch, T.; Scherer, K.R.; Grandjean, D.M.; Sander, D. The impact of emotion on perception, attention, memory, and decision-making. *Swiss Med. Wkly.* **2013**, *143*, w13786. [CrossRef]
13. Fleck, S.J.; Kraemer, W. *Designing Resistance Training Programs*, 4th ed.; Human Kinetics: Leeds, UK, 2014.
14. Pallavicini, F.; Argenton, L.; Toniazzi, N.; Aceti, L.; Mantovani, F. Virtual reality applications for stress management training in the military. *Aerosp. Med. Hum. Perform.* **2016**, *87*, 1021–1030. [CrossRef]
15. Ćosić, K.; Popović, S.; Horvat, M.; Kukolja, D.; Dropuljić, B.; Kostović, I.; Judas, M.; Rados, M.; Rados, M.; Vasung, L.; et al. Virtual reality adaptive stimulation in stress resistance training. In Proceedings of the RTO-MP-HFM-205 on Mental Health and Well-Being across the Military Spectrum, Bergen, Norway, 11–13 April 2011.
16. Kothgassner, O.D.; Goreis, A.; Kafka, J.X.; Van Eickels, R.L.; Plener, P.L.; Felnhofer, A. Virtual reality exposure therapy for posttraumatic stress disorder (PTSD): A meta-analysis. *Eur. J. Psychotraumatol.* **2019**, *10*, 1654782. [CrossRef]
17. Deng, W.; Hu, D.; Xu, S.; Liu, X.; Zhao, J.; Chen, Q.; Liu, J.; Zhang, Z.; Jiang, W.; Ma, L.; et al. The efficacy of virtual reality exposure therapy for PTSD symptoms: A systematic review and meta-analysis. *J. Affect. Disord.* **2019**, *257*, 698–709. [CrossRef]
18. Horvat, M.; Dobrinić, M.; Novosel, M.; Jerčić, P. Assessing emotional responses induced in virtual reality using a consumer EEG headset: A preliminary report. In Proceedings of the 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO 2018), Opatija, Croatia, 21–25 May 2018; IEEE: Piscataway Township, NJ, USA, 2018; pp. 1006–1010.
19. Yang, K.; Wang, C.; Gu, Y.; Sarsenbayeva, Z.; Tag, B.; Dingler, T.; Wadley, G.; Goncalves, J. Behavioral and physiological signals-based deep multimodal approach for mobile emotion recognition. *IEEE Trans. Affect. Comput.* **2021**. [CrossRef]
20. Doradzińska, Ł.; Furtak, M.; Bola, M. Perception of semantic relations in scenes: A registered report study of attention hold. *Conscious. Cogn.* **2022**, *100*, 103315. [CrossRef]
21. Kotowski, K.; Stapor, K. Machine Learning and EEG for Emotional State Estimation. *Sci. Emot. Intell.* **2021**, 75.

22. Marchewka, A.; Żurawski, Ł.; Jednorog, K.; Grabowska, A. The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behav. Res. Methods* **2014**, *46*, 596–610. [CrossRef]
23. Lang, P.J.; Bradley, M.M.; Cuthbert, B.N. *International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual*; Technical Report A-8; University of Florida: Gainesville, FL, USA, 2008.
24. Dan-Glauser, E.S.; Scherer, K.R. The Geneva affective picture database (GAPED): A new 730-picture database focusing on valence and normative significance. *Behav. Res. Methods* **2011**, *43*, 468–477. [CrossRef]
25. Kurdi, B.; Lozano, S.; Banaji, M.R. Introducing the open affective standardized image set (OASIS). *Behav. Res. Methods* **2017**, *49*, 457–470. [CrossRef]
26. Haberkamp, A.; Glombiewski, J.A.; Schmidt, F.; Barke, A. The DIsgust-RelaTed-Images (DIRTI) database: Validation of a novel standardized set of disgust pictures. *Behav. Res. Ther.* **2017**, *89*, 86–94. [CrossRef]
27. Posner, J.; Russell, J.A.; Peterson, B.S. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Dev. Psychopathol.* **2005**, *17*, 715. [CrossRef]
28. Bakker, I.; Van Der Voordt, T.; Vink, P.; De Boon, J. Pleasure, arousal, dominance: Mehrabian and Russell revisited. *Curr. Psychol.* **2014**, *33*, 405–421. [CrossRef]
29. Evmenenko, A.; Teixeira, D.S. The circumplex model of affect in physical activity contexts: A systematic review. *Int. J. Sport Exerc. Psychol.* **2022**, *20*, 168–201. [CrossRef]
30. Atmaja, B.T.; Sasou, A.; Akagi, M. Survey on bimodal speech emotion recognition from acoustic and linguistic information fusion. *Speech Commun.* **2022**, *140*, 11–28. [CrossRef]
31. Von Scheve, C. Emotion regulation and emotion work: Two sides of the same coin? *Front. Psychol.* **2012**, *3*, 496. [CrossRef] [PubMed]
32. Susskind, J.M.; Lee, D.H.; Cusi, A.; Feiman, R.; Grabski, W.; Anderson, A.K. Expressing fear enhances sensory acquisition. *Nat. Neurosci.* **2008**, *11*, 843–850. [CrossRef]
33. Burkitt, E.; Watling, D.; Cocks, F. Mixed emotion experiences for self or another person in adolescence. *J. Adolesc.* **2019**, *75*, 63–72. [CrossRef]
34. Tottenham, N.; Tanaka, J.W.; Leon, A.C.; McCarry, T.; Nurse, M.; Hare, T.A.; Marcus, D.J.; Westerlund, A.; Casey, B.J.; Nelson, C. The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Res.* **2009**, *168*, 242–249. [CrossRef]
35. Dawel, A.; Miller, E.J.; Horsburgh, A.; Ford, P. A systematic survey of face stimuli used in psychological research 2000–2020. *Behav. Res. Methods* **2021**, 1–13. [CrossRef]
36. Stevenson, R.A.; Mikels, J.A.; James, T.W. Characterization of the affective norms for English words by discrete emotional categories. *Behav. Res. Methods* **2007**, *39*, 1020–1024. [CrossRef]
37. Olszanowski, M.; Pochwatko, G.; Kuklinski, K.; Scibor-Rylski, M.; Lewinski, P.; Ohme, R.K. Warsaw set of emotional facial expression pictures: A validation study of facial display photographs. *Front. Psychol.* **2015**, *5*, 1516. [CrossRef]
38. Stevenson, R.A.; James, T.W. Affective auditory stimuli: Characterization of the International Affective Digitized Sounds (IADS) by discrete emotional categories. *Behav. Res. Methods* **2008**, *40*, 315–321. [CrossRef]
39. Horvat, M.; Jović, A.; Burnik, K. Assessing the Robustness of Cluster Solutions in Emotionally-Annotated Pictures Using Monte-Carlo Simulation Stabilized K-Means Algorithm. *Mach. Learn. Knowl. Extr.* **2021**, *3*, 435–452. [CrossRef]
40. Trnka, M.; Darjaa, S.; Ritomský, M.; Sabo, R.; Rusko, M.; Schaper, M.; Stelkens-Kobsch, T.H. Mapping Discrete Emotions in the Dimensional Space: An Acoustic Approach. *Electronics* **2021**, *10*, 2950. [CrossRef]
41. Bynion, T.M.; Feldner, M.T. Self-assessment manikin. *Encycl. Personal. Individ. Differ.* **2020**, 4654–4656. [CrossRef]
42. Morris, J.D. Observations: SAM: The Self-Assessment Manikin; an efficient cross-cultural measurement of emotional response. *J. Advert. Res.* **1995**, *35*, 63–68.
43. Wierzba, M.; Riegel, M.; Pucz, A.; Leśniewska, Z.; Dragan, W.Ł.; Gola, M.; Jednorog, K.; Marchewka, A. Erotic subset for the Nencki Affective Picture System (NAPS ERO): Cross-sexual comparison study. *Front. Psychol.* **2015**, *6*, 1336. [CrossRef]
44. Zamora, E.V.; Richard's, M.M.; Introzzi, I.; Aydmune, Y.; Urquijo, S.; Olmos, J.G.; Marchewka, A. The Nencki Affective Picture System (NAPS): A Children-Rated Subset. *Trends Psychol.* **2020**, *28*, 477–493. [CrossRef]
45. Constantinescu, A.C.; Wolters, M.; Moore, A.; MacPherson, S.E. A cluster-based approach to selecting representative stimuli from the International Affective Picture System (IAPS) database. *Behav. Res. Methods* **2017**, *49*, 896–912. [CrossRef]
46. Stappen, L.; Schumann, L.; Sertolli, B.; Baird, A.; Weigell, B.; Cambria, E.; Schuller, B.W. Muse-toolbox: The multimodal sentiment analysis continuous annotation fusion and discrete class transformation toolbox. In Proceedings of the 2nd on Multimodal Sentiment Analysis Challenge, Virtual Event, China, 24 October 2021; pp. 75–82.
47. Azevedo, I.L.; Keniston, L.; Rocha, H.R.; Imbiriba, L.A.; Saunier, G.; Nogueira-Campos, A.A. Emotional categorization of objects: A novel clustering approach and the effect of depression. *Behav. Brain Res.* **2021**, *406*, 113223. [CrossRef]
48. Redies, C.; Grebenkina, M.; Mohseni, M.; Kaduhm, A.; Dobel, C. Global image properties predict ratings of affective pictures. *Front. Psychol.* **2020**, *11*, 953. [CrossRef]
49. Hoffmann, H.; Scheck, A.; Schuster, T.; Walter, S.; Limbrecht, K.; Traue, H.C.; Kessler, H. Mapping discrete emotions into the dimensional space: An empirical approach. In Proceedings of the 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Seoul, Korea, 14–17 October 2012; pp. 3316–3320.

50. Kervadec, C.; Vielzeuf, V.; Pateux, S.; Lechervy, A.; Jurie, F. Cake: Compact and accurate k-dimensional representation of emotion. *arXiv* **2018**, arXiv:1807.11215.

51. Bojanić, M.; Gnjatović, M.; Sečujski, M.; Delić, V. Application of dimensional emotion model in automatic emotional speech recognition. In Proceedings of the 2013 IEEE 11th International Symposium on Intelligent Systems and Informatics (SISY), Subotica, Serbia, 26–58 September 2013; pp. 353–356.

52. Ricci, F.; Rokach, L.; Shapira, B. (Eds.) Recommender systems: Techniques, applications, and challenges. In *Recommender Systems Handbook*; Springer: New York, NY, USA, 2022. [CrossRef]

53. Deldjoo, Y.; Schedl, M.; Cremonesi, P.; Pasi, G. Recommender systems leveraging multimedia content. *ACM Comput. Surv. (CSUR)* **2020**, *53*, 1–38. [CrossRef]

54. Raheem, K.R.; Ali, I.H. Survey: Affective recommender systems techniques. In Proceedings of the IOP Conference Series: Materials Science and Engineering, Thi-Qar, Iraq, 15–16 July 2020; Volume 928, p. 032042.

55. Santamaria-Granados, L.; Munoz-Organero, M.; Ramirez-Gonzalez, G.; Abdulhay, E.; Arunkumar, N.J.I.A. Using deep convolutional neural network for emotion detection on a physiological signals dataset (AMIGOS). *IEEE Access* **2018**, *7*, 57–67. [CrossRef]

56. Salazar, C.; Aguilar, J.; Monsalve-Pulido, J.; Montoya, E. Affective recommender systems in the educational field. A systematic literature review. *Comput. Sci. Rev.* **2021**, *40*, 100377. [CrossRef]

57. Diaz, Y.; Alm, C.O.; Nwogu, I.; Bailey, R. Towards an affective video recommendation system. In Proceedings of the 2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), Athens, Greece, 19–23 March 2018; pp. 137–142.

58. Elahi, M.; Ricci, F.; Rubens, N. A survey of active learning in collaborative filtering recommender systems. *Comput. Sci. Rev.* **2016**, *20*, 29–50. [CrossRef]

59. Kroese, D.P.; Brereton, T.; Taimre, T.; Botev, Z.I. Why the Monte Carlo method is so important today. *Wiley Interdiscip. Rev. Comput. Stat.* **2014**, *6*, 386–392. [CrossRef]

60. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65. [CrossRef]

61. White, C.N.; Liebman, E.; Stone, P. Decision mechanisms underlying mood-congruent emotional classification. *Cogn. Emot.* **2018**, *32*, 249–258. [CrossRef]

62. Horvat, M.; Jednoróg, K.; Marchewka, A. Clustering of affective dimensions in pictures: An exploratory analysis of the NAPS database. In Proceedings of the 39th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO 2016), Opatija, Croatia, 30 May–3 June 2016; pp. 1496–1501.

63. Horvat, M.; Popović, S.; Ćosić, K. Towards semantic and affective coupling in emotionally annotated databases. In Proceedings of the 35th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO 2012), Opatija, Croatia, 21–25 May 2012; pp. 1003–1008.