A Case-Based Approach to Creating Movie Poster Compositions

Jorge Alejandro Ramírez Gallardo
Mechatronic and Industrial Engineering
Instituto Tecnológico Autónomo de México (ITAM)
Mexico City, Mexico
jorge.ramirez.gallardo@itam.mx

Brandon Francisco Hernández Troncoso
Mechatronic and Telecommunications Engineering
Instituto Tecnológico Autónomo de México (ITAM)
Mexico City, Mexico
brandon.hernandez@itam.mx

Rodrigo Castillo Gómez
Computer and Mechatronic Engineering
Instituto Tecnológico Autónomo de México (ITAM)
Mexico City, Mexico
rodrigocastillog97@gmail.com

Andrés Gómez de Silva Garza
Department of Computer Engineering
Instituto Tecnológico Autónomo de México (ITAM)
Mexico City, Mexico
agomez@itam.mx

Abstract

This paper describes a system for the design or redesign of movie posters. The main reasoning strategy used by the system is case-based reasoning (Leake 1996), which requires two important sub-tasks to be performed: case retrieval and case adaptation. We have used the random forest algorithm to implement the case retrieval subtask, as it allows the system to formulate a generalization of the features of all the top matching cases that are determined to be most relevant to the new (re)design desired. We have used heuristic rules to implement the case adaptation task which results in generating the suggestion for poster composition. These heuristic rules are based on the Gestalt theory of composition (Arnheim 1974) and the requirements specified by the user.

Introduction

The main objective of our system is to design new posters for movies based on certain desired characteristics. On the other hand, the system can also be used to redesign an already existing poster for a movie. The system's name is CACARO, which stands for CAse Composition Algorithm for the (Re)design of One-Sheets. The term "one-sheet" is used in the movie industry to refer to a movie poster. The generic name "Cácaro" is used in Mexico to refer to the projectionist in a movie theater—anyone who is a projectionist is called "Cácaro" by the paying public.

The user describes a new problem to the system by specifying the most important desired characteristics and CACARO fills out the description of the poster by proposing additional characteristics. These additional characteristics are decided based on already known (previously existing) posters within the system's case base (which contains 103 cases).

In order to transfer the knowledge from past cases to the new situation, the abstraction and generation of knowledge required to describe the new poster is based on the Gestalt theory of composition.

The structure of the rest of this paper is as follows. First, we provide a section in which we discuss the theory of composition in the context of poster composition design. Then we present sections on decision trees and random forests, including how we use them for case retrieval in our system, and on heuristic rules and how we use them for case adaptation in our system. We then include a section that provides implementation details. Additionally, we have a section in which we fully trace, and describe the results, of running one test problem on our system and include a brief discussion of the results from two additional test problems. Finally, we provide a section which includes discussions, conclusions, and future work.

Theory of Composition

To propose the composition of a movie poster, CACARO uses Gestalt theory and its principles of perceptual organization. The main approach of Gestalt principles to art is based on Arnheim's vision (1974), which recognizes that the whole (in this case, the poster) is much more than just a collection of its parts (in this case, characters, background colors or decorations, and objects/props). Previous work (Desolneux, Moisan and Morel 2008; Guberman 2015; Kobourov, Mchedlidze and Vonessen 2015) has implemented Gestalt principles for computer vision, graph drawing and image analysis, but our system uses them to create the poster's composition.

To achieve a good composition, the poster should be in equilibrium. The main reference for equilibrium in composition is balance, which establishes that each figure or element inside the composition has come to a standstill and "no change seems possible" (Arnheim 1974). The two main properties that affect the balance are weight and direction.

Weight refers to a visual element's capacity to generate tension in the composition and it is defined by taking into account some of the characteristics of the element. Some quantifiable characteristics are location (an element is considered heavier if it is further away from a poster's center), spatial depth (an element is considered lighter if it looks closer than those that are seemingly further from the viewer), size, color (an element is considered heavier if it is brighter or more reddish), and isolation (Arnheim 1974).

However, other characteristics of a visual element like intrinsic interest (spectator preferences) and shape (how the viewer perceives object boundaries and their axes or structural skeletons) are related to human perception (Arnheim 1974), and therefore cannot be as easy to quantify. In our framework, this potential drawback is mitigated by the use of Case-Based Reasoning (CBR), which encodes in the cases these qualitative human "measurements" (opinions) and combines both the quantitative and qualitative descriptors of known posters in order to propose the composition of new posters.

The main purpose for using CBR is to allow the reuse of old solutions, in this case existing movie posters, to meet new demands, in this case new movie posters we want to generate (Kolodner 1993). In other words, using CBR allows us to retrieve and compare features of the posters stored in the knowledge base to generate a new poster following the new requirements and the Gestalt composition theory. The reason why this is useful is because in art and design it is common to use inspiration from previous and similar artifacts rather than beginning the design of a new one from scratch.

The second property related to balance is *direction*, which refers to the possibility of a viewer perceiving some visual elements in a composition as pointing somewhere (perhaps towards another object), creating visual lines that result from an element's neighboring elements, subject matter, and shape (Arnheim 1974). Only the first trait is easily quantifiable, but again, CBR helps to combine these with qualitative attributes.

Balance is reachable through many strategies (Arnheim 1974) that assign weight differently in order to obtain a stable mass center. Some of them include an analysis of the symmetry axis (vertical, horizontal, or diagonal), top vs. bottom balance (where top elements are heavier or lighter than bottom elements of the same size and location) and right vs. left balance (which applies if no symmetrical balance exists because elements on the right are heavier than elements on the left, or vice versa). The calculation of mass centers (McManus, Stöver and Kim 2011), weights, distances, and correlations between objects (like similarity or overlap of color positioning) is necessary in all the mentioned strategies.

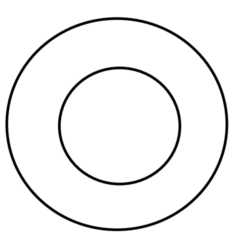


Figure 1. Illustration of a disc.

Figure 2. Two types of football: on the left a soccer ball and on the right an American football.

The final aspect of a good composition is a defined form for each figure. The *form* refers to the viewer's perception of an object regarding composition, which is not necessarily the same as the shape of the object (Arnheim 1974). Consider Figure 1, where a disc is displayed. Although the image shape is that of a disc, without this information a person could infer that the image shows two circumscribed circles, a tire, or even a donut. Thus, composition should provide enough details of each visual element to communicate its intended interpretation as clearly as possible.

Because CACARO suggests a composition for a poster, each visual element's form can only be altered by *foreshortening* or *overlapping*. These two strategies must take into account image aspect ratio, framing, and continuity, so the element should not look distorted or amputated (Arnheim 1974). Additionally, it is convenient in posters to consider the concepts of *positive and negative spaces*, because the blank areas (negative space) that appear together with the objects of interest (positive space) are a "critical composition element", as Suler (2013) points out.

Decision Trees and Random Forests

In machine learning, classification problems can be described using *decision trees*. Each internal node of a decision tree contains a question (or, at its most basic, the name of a decision variable), and the set of branches departing that node represent the different potential answers to the question (or values for the variable). By answering the questions and following the branches with respect to a given instance (example) which one wants to classify, each time gaining more distance from the root node, the search becomes more and more focused, and the search space more and more reduced. Each leaf node represents a possible classification (decision, prediction, value) for the instance which is being analyzed using the decision tree.

It is very important to highlight that many different trees can be built to solve a single problem. Figure 2 shows two different kinds of "footballs": on the left, the classic soccer ball, round with interlocking pentagonal and hexagonal shapes, usually black and white, etc., and on the right an American (gridiron, NFL) football, shaped like a pointy oval with tiny bumps on the surface, usually brown, etc.

Let us assume that a system is trying to classify a new instance that has been thrown at it as either a soccer ball (+) or an American football (-). Depending on the features we decide to analyze, the ambiguity (or lack thereof) in the values of these attributes, and even which of these parameters we use in the root node of a classification tree, we may get very different trees with the same aim, as Figure 3 illustrates.

In Figure 3 one can observe that a big advantage of decision tree classifiers is interpretability. A human being can see the decision tree and understand the reasoning behind the classifications, which is not the case with many other machine learning classifiers.

When creating a decision tree, an important question is which question or variable should be included at each level of the tree. Two of the most popular options are the use of *information gain* and the use of the *Gini index* to make the decision. Both metrics aim to build compact decision trees, since the larger a tree gets, the more it tends to under-generalize (overfit) the training data. Another approach to avoid overfitting is the use of pruning techniques which reduce the tree's size. This might lead to lower accuracy in the tree's description of the training data, but it leads to better generalization and therefore better results when the tree is used to classify unknown data (Shalev-Shwartz and Ben-David 2014; Kubat 2017).

However, in CACARO, to reduce the risk of overfitting we used *random forests*. A random forest is a collection of different decision trees applied to the same problem. Each tree classifies an instance individually and then, through majority voting, a final classification is proposed. The trees inside a forest are all different from each other and use only a subset of parameters to make their classification. The generalization error of a forest depends on the strength of its trees and the correlation between them. Using random subsets of attributes to do this not only reduces the error but increases the classifier's tolerance to noise (Breiman 2001).

CACARO runs the random forest algorithm five times on the same problem description with different sets of hyper-parameters in each of these runs. Some such hyper-parameters used in the random forest algorithm are the number of features that are analyzed, the weights of the different features that are considered and using different random seeds for each forest. The result is a list of up to five different cases that are deemed to be most relevant by the random forest algorithm (though there could be duplicates despite varying the values of the hyper-parameters). Despite the various cases that are retrieved, they are all relevant to the problem description or they would not have been the random forest algorithm's selection. This method for case retrieval is not unique to our work. It was analyzed at length along with other alternative case-retrieval techniques in (Löw et al. 2019). In addition, this technique was found to increase the effectiveness of case retrieval in (Yang and Wu 2001). Still, even though it is not the most widely used case retrieval method, given our application domain we found it appropriate.

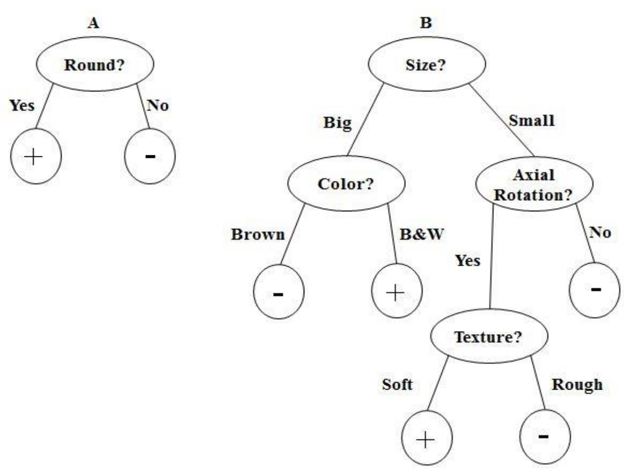


Figure 3. Two alternative decision trees for the same problem (classifying an object as a soccer ball or an American football).

Once the system retrieves relevant cases through the use of the random forests, we need a way to use the knowledge implicitly held in them to propose a description that contains the main characteristics of these relevant cases represented in one general description for the new poster composition. To accomplish this case adaptation, the system uses heuristic rules based on the theory of composition.

Heuristic Rules

After relevant cases have been recovered, the system proposes values for attributes that were not included in the initial problem specification by abstracting a generic case from the recovered cases through the use of heuristic rules. *Heuristic rules* are used to find solutions for problems in artificial intelligence. These rules trade accuracy and completeness in favor of speed and performance in their decision making (Pearl 1984). These rules are often used in bundles to break up complex problems into more simple ones and to find approximations to the solution by aggregation (Romanycia and Pelletier 1985).

CACARO breaks up the problem of synthesizing a resulting proposal from the outputs of our random forests into several heuristic rules which adjust select features from the representations of retrieved cases. This results in obtaining a case adaptation which synthesizes case properties, thus generating better poster proposals than if the knowledge held in only one retrieved case were used. In other words, we consider that all matching cases might be mined for useful information in proposing the new composition, not just the top matching case.

Implementation

The algorithms that process a new problem and implement the case retrieval and case adaptation sub-tasks in CACARO are written in Python. This programming language provides functionality that makes it easy to process large amounts of data without writing large amounts of code, and also has many predefined libraries that simplify the processing of visual information.

Each poster in our system is represented as a separate case. Some of the information stored in each case includes values for attributes such as the number of characters from the movie shown in the poster, the type and number of objects/props that are present in the poster, the type of background included in the poster, and so on. Each case is described using the same standard set of attributes, which allows for easy comparisons among the cases and easy probing of case memory to determine the cases that are relevant given the description of a new desired poster.

Figure 4 shows the general description and structure of a case stored in CACARO. Each case can be described by at most 29 attributes, including both the description of the solution and the problem requirements for a given case. The attributes used for each case are classified by Category and Type of Information.

Categories in CACARO are groups of features related to the description of a poster element. There are four categories: Description (general information about the case), Composition (elements considered for the composition of the poster), Characters (description, image and properties of each character shown in the poster), and Visualization (characteristics of the poster that affect its visual perception which are not included in any of the previous categories). Each attribute in Figure 4 is prefixed by the initial of its category name (in the case of Characters, their attributes are prefixed by a “CH” rather than just the initial, to distinguish from Composition attributes, which are prefixed by a "C"). For example, Movie Name is an attribute belonging to the Description category, so it is prefixed by a “D”.

Consideration for *Type of Information* classifies the features of a case based on their role in the poster design process. The categories here are: Goal, Input Image Requirement (IIR), and Outcome. Goals are the user requirements for a poster and they contain the information that defines the main properties required in a new poster (new case). On the other hand, for the cases initially contained in CACARO’s case base (old cases), we manually assigned the values for the attributes classified as Goals based on the perceived characteristics and final appearance of the corresponding poster. This is due to the fact that the initial cases were not the solutions to problems posed by the user.

For the design of a new case (poster) , values for attributes described as Goals must be provided by the user, except for the attributes marked with an asterisk, which are optional for

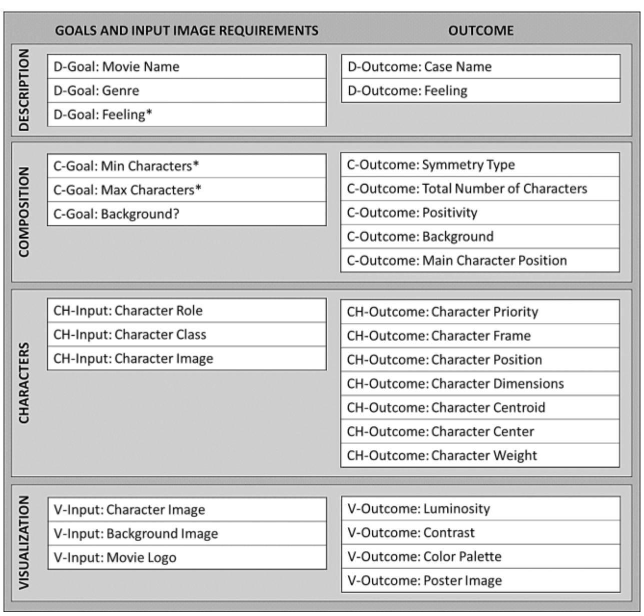


Figure 4. Structure of a general case in CACARO.

the user to provide (for example, Feeling) if more restrictions than the minimum amount need to be specified.

An IIR attribute describes the information provided by the user which identifies and includes the images inserted for a new case. These images must be about the characters, backgrounds (scenery, locations), and/or the logo that can appear in the new poster.

Finally, attributes described as Outcome contain the descriptions of the case which are the result of the composition and design processes performed by CACARO. In other words, the role of CACARO is to fill in the values of some or all of these attributes as a result of the decision-making process that the system performs in order to solve a new poster composition problem.

Some of the attributes, like positivity, are the result of performing an analysis of the image in the poster and its goal and input requirement attributes. Therefore, the solution part of each case within the case base is fully described using only 15 stored attributes: *Case Name, Movie Name, Genre, Feeling, Symmetry Type, Total Number of Characters, Color Palette, Background Image, Movie Logo, Character (Role, Class, Image, Priority, Framing)* and *Poster Image*. Each attribute in the case base has a specific type of value. Valid values for each attribute depend on boundaries defined by the associated data type and domain knowledge.

Some of the attributes contain the data directly. For example, features like Movie Name contain string data (without allowing null values). Character Image, Movie Logo, and Background Image are PNG images. Some attributes described by integer data are Luminosity and Contrast (derived attributes), with valid values between 0 and 255.

However, most of the attributes were represented internally using bit strings, a process which allows them to be easily used in the random forest and the heuristic functions used in CACARO (which are programmed to manipulate bit strings). The encoding techniques used to represent those attributes were the one-hot encoding and binary encoding techniques, depending on the attribute. It is important to mention that binary encoding allows the description of each exemplar through one or more categories (values) at the same time, whereas one-hot encoding is intended to be used when categories are mutually exclusive (Cohen et al. 2013).

The binary encoding technique is used to represent c*ategorical data* in a binary format, where each category is associated to a position within a bit string and the bit is turned

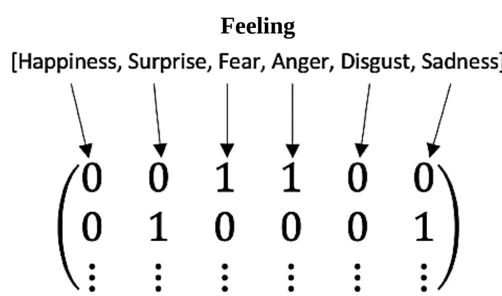


Figure 5. Example of the binary encoding technique in CACARO.

on or off depending on the presence or absence of the descriptive category in a particular example. Figure 5 is an example of the binary encoding technique applied to the *Feeling* attribute (the only attribute which uses this encoding scheme in CACARO).

In Figure 5, each row shows the description of one movie poster, and both examples shown use two sentiments to describe the feeling or mood that the poster design was intended to convey (fear and anger in the case of the first movie, and surprise and sadness in the case of the second). However, a movie poster can be associated to only one sentiment or to more than two.

On the other hand, the one-hot encoding technique similarly assigns one bit position within a bit string to each possible value of a category but limits the entire bit string to only have one bit turned on (Pai, Pardawala and Potdar 2017; Harris and Harris 2013). Figure 6 shows an example of this technique applied to the Genre attribute. Again, the figure shows the descriptions of two movies, one per row. One of the movies is classified as a Romance film and the other as a Horror movie.

Each case in CACARO was stored in its own row within an Excel file stored in .csv format which facilitates loading them into the random forest code that uses *scikitlearn*. This Excel file represents the system's case base. The features taken for this file are: *Poster Name (file), Movie Name, Genre, Feeling, Symmetry, Number of Characters, Color Pallete, Character1 Class, Character1 Priority, Character1 Role, Character1 Framing* (and the same for characters 2 and 3 if they're needed), *Background Description, Background Priority, Object Description,* and *Object Priority.* It bears mentioning that each "Priority" feature is a reference to the importance of the character in relation to the others,

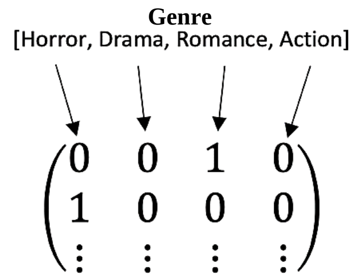


Figure 6. Example of the one-hot encoding technique in CACARO.

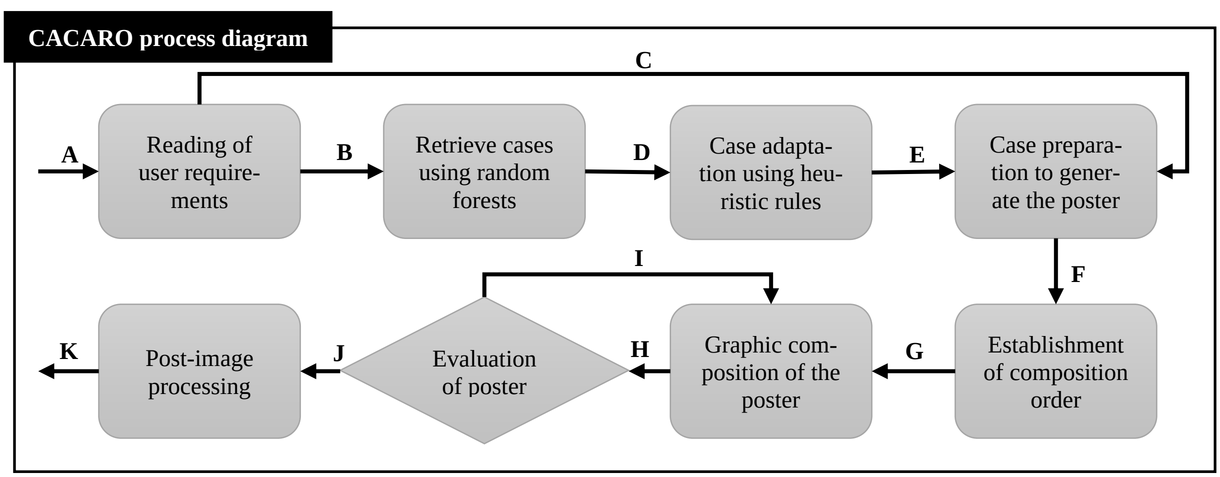
affecting their relative size within the poster. Also, background and object description are used to mean the presence or absence of either within a given case.

Given the fact that random forests only compare numerical values, most of the values were codified with one-hot encoding. An example of the codification proposed for one of the features, *Genre*, is as follows:

Genre
Terror = 0001
Drama = 0010
Romance = 0100
Action = 1000

If there is no object or background then the related features in the Excel file will be assigned a value of −10, and the same occurs with characters 2 and 3 if they're absent from a given poster.

In order to explain CACARO's workflow, Figure 7 shows all the main blocks and step-by-step tasks that CACARO



A. User request for a new poster.
B. Vector of poster characteristics defined by the user.
C. Character data inserted by the user (character class and image).
D. Five retrieved cases.
E. One final case resulting from adaptation.
F. Information required for composition order.
G. Information required for poster composition.
H. Proposed poster composition.
I. Poster does not fit quality requirements.
J. Poster fits with quality requirements.
K. Final version of the poster.

Figure 7. CACARO's process diagram.

performs, which we describe in detail in the following paragraphs. Each input and output of each process is listed below the diagram and referenced in it with a capital letter. In the beginning, the user inputs a request to generate a new poster by giving some characteristics of the desired poster, for example the genre of the movie, a color palette, the number of the characters and some images of those characters. Once the input with the features have been introduced, CACARO translates those features into a special vector to be used as input in the random forests.

Random forests are instantiated and trained using the *scikitlearn* module for Python. Given the fact that each poster is a different and unique case in the case base, they can't be grouped together to classify a new instance. Because of this fact we made the decision that the individual posters are leaf nodes in each of the decision trees.

For each design problem the system produces five random forests with different hyper-parameters (for example, different amounts of features per tree, different random seeds which allow the experiments to be reproducible and the random forests to be different from one another, and also different number of trees per forest). The leaf nodes of the resulting forests are posters which contain some of the characteristics that the user desires (included in the initial problem specification). After gathering the five results, only the set of cases corresponding to these five posters is recovered from the case base. CACARO proposes values for attributes that were not included in the initial problem specification by abstracting a generic case from the recovered cases through the use of several heuristic rules that synthesize the data held in the five retrieved cases. The following paragraph explains some of these synthesis rules.

For the number of characters, we get a range based on the average of number of characters found in the recovered cases. For the symmetry, we randomly obtain a symmetry from those in the recovered cases. For the poster's colors, we take the top five colors present in each case (determined by measuring and comparing the total number of pixels occupied by each color) and find the mean saturation. We then do the same for each of the other four cases, and finally find the mean of the five means. An analogous process is followed in order to determine the hue. Both the new saturation and the new hue found through this process are used as a filter when deciding the color of the resulting poster(s). For the positivity, an attribute whose semantics we describe in more detail below, we obtain the median of the recovered cases and establish a range of ±12% from this value as the acceptable range of values for the proposed cases.

We think that these heuristics which combine some random decisions within limits obtained from the cases in the case base serve to give CACARO its potential for creativity. CACARO has a second set of heuristics which yield recommendations that depend on the inputs the user gave to initiate the (re)design process. These heuristics refer to the number and type of characters and their framings (the percentage of each character that appears in the poster—only face, full body, etc.—depending on the character image(s) provided by the user).

All the previous heuristic rules determine values for each of the relevant parameters in the description of a poster's composition. These values are used in the case preparation phase to generate the resulting poster. In this phase, CACARO makes use of the information related to the heuristics and also the user information about characters. It establishes the initial position and size of the images within the poster it will propose and it also uses the character class specified by the user later, for the evaluation of the proposed composition of the poster.

To perform the poster composition, we need to consider the following premises based on the knowledge base and the visual posters considered: 1) the poster must have a background; 2) the characters do not have any specific predefined order in the composition; and 3) the size of the characters in the poster will be defined by their role and importance in the actual movie. With these premises in mind, the order of the composition is performed randomly for each proposed poster, changing the size of each character and their position within the poster. Using this technique, CACARO can find the best composition by evaluating it using the heuristic rules described above.

To assess the quality of the composition of a proposed poster, some characteristics of the image are evaluated. If one or more characteristics do not fulfill the requirements or thresholds established by the heuristics, CACARO performs the actions needed to improve the composition and meet the thresholds. These characteristics are evaluated by four rules based on Gestalt theory.

The first rule evaluates the *proportions criterion*, which refers to how much the character image can be resized and moved within the poster; therefore, it is not a criterion with degrees of compliance, but instead a hard requirement that the image must accomplish to be considered a good poster composition according to the evaluation of the relations between composition elements as described by Arnheim (1974).

CACARO compares the relative sizes of the different characters included in the poster and the ratio of the character's size to that of the poster itself. It also determines if those ratios fulfill the criterion by comparing them with desired intervals. Each desired interval is different based on characters' classes and priorities, and their values were determined empirically based on the range displayed by the relevant posters in the case base. For instance, if CACARO compares within a poster proposal a character X (class=protagonist, priority=2) with a character Y (class=secondary, priority=3), the proportions criterion will accept the size ratio of X to Y only if its value falls between 0.7 and 0.9 because those are the limits of the size ratios between protagonists and secondary characters with respective priorities of 2 and 3 within the posters that are included in the case base.

The second evaluation, related to the posters' *positivity*, is based on the degree of overlap and sizing of the characters in the posters. This is done by using a pixel matrix (*positivity matrix*) the same size as the poster in which each pixel occupied by a character is set to 1 and the rest is set to 0. Positivity is equal to the number of pixels occupied by a character divided by the total number of pixels in the poster. As explained in the heuristics, this rule takes into account the positivity of the recovered cases and obtains the average positivity, which is used to define the acceptable range.

The third rule stems from the concept of *balance* from Gestalt theory. For its evaluation, CACARO uses the selected axis of symmetry for the current composition to split the positivity matrix. This matrix shows only the position and form of the characters, which are the equivalent of the tension elements described in Gestalt theory (Arnheim 1974). But, in order to calculate balance, we need to determine the weight of each element given its position, as Gestalt theory specifies (McManus, Stöver and Kim 2011). Accordingly, we defined a *weight matrix* which represents the distribution of the visual tension described by Arnheim (1974), where the top-right pixel (TRP) is the heaviest part and the bottom-left pixel (BLP) is the lightest. This matrix is used for all compositions and its values are defined as 0.1 for BLP and 1.0 for TRP.

The weight matrix is multiplied by the split positivity matrix to quantify the tension generated by the two halves of the poster. Then, the tension difference (in percentage) is calculated. If its value is less than 10% the poster is considered as balanced.

Lastly, the system performs *face detection* of the characters in the poster to evaluate whether the composition of the characters is correct or a resize or translation is needed. If the faces of two characters overlap by at least 20% of the rear face, one character is moved to eliminate the excessive overlap. CACARO accomplishes this by implementing YOLO's neural network architecture (Redmon and Farhadi 2018), trained with 600 hand-labeled images with faces in them, taken from the WIDER FACE dataset (Yang, Luo, Loy, and Tang 2016).

Regarding the rules that CACARO uses to evaluate the posters, each one assigns a degree of quality according to a specific criterion. These measurements are averaged to obtain the overall quality of the poster, which can be assigned from 0 to 100%. In this way we can establish a minimum quality threshold needed to accept a composition for the poster as a valid solution.

Regarding the *proportions criterion*, the degree of quality is measured according to the percentage of size ratios that fall within the acceptable interval of values. We set CACARO's threshold for this criterion at 100%, which means that all the size ratios must be within the corresponding interval in order to achieve an acceptable quality solution.

The *positivity rule* guarantees the poster's positivity is within the acceptable range set by the rule. We consider the difference between the average positivity and the proposed poster's positivity to define the degree of quality of the poster according to this criterion by applying the following formula:

$$Positivity\ quality = \left\{1 - \left|\frac{poster's\ positivity - \ positivity\ average}{positivity\ average}\right|\right\} \times 100\%$$

We set CACARO's threshold at a minimum of 88% to ensure quality but also to allow variability and creativity in the generated posters.

The measurement of the degree of quality for the last two criteria (*balance* and *face detection*) is similar to the previous one, but instead of the difference in positivities, *balance* uses the difference in tensions and *face detection* uses the percentage of overlap between faces. We set their threshold levels in CACARO at 90% and 80%, respectively.

As explained, the four criteria need to be met so a proposed solution can be considered "acceptable". In order to fully converge CACARO first evaluates each criterion in the order presented in this paper. If a criterion was not met, CACARO makes the previously explained changes and again evaluates each criterion. If any of the previously met criteria is not met anymore due to the modifications, CACARO will start making further modifications, and reevaluating, until all criteria are met.

As a result of the previous implementation phases, CACARO provides one resulting poster which must fulfill the user requirements and must be generated by evaluating the metrics and meeting the criteria described above. CACARO performs the poster composition process based only on Gestalt theory (which might not be in accordance with criteria from other design theories).

However, CACARO also generates several additional alternative compositions for the poster. This allows the user to perform a visual comparison to complement the algorithmic decisions made by the system previously. The user can then decide which is the best alternative. This framework can also improve creativity through the variety of options present among the proposals. A saturation and value filter is applied to the final proposals in order to get similar saturations and values to the retrieved cases.

Experiments and Results

We tested CACARO with several problems to evaluate the quality of the results. In this section, however, we only present three of these experiments, one of them fully detailed and the other two just presented through their inputs and outputs.

In the first test problem, the goal was the redesign of the poster for "The Prestige" (Touchstone Pictures 2006). The most important input requirements were "Drama" (as *Genre*) and the images and features (*Class* and *Role*) of the three main characters of the movie. Based on all the inputs, the random forests suggested that five most relevant cases are "Men Of Honor" (Twentieth Century Fox 2000), "The Shawshank Redemption" (Columbia Pictures 1994), and "Kick-Ass 2" (Universal Pictures 2013), the last one retrieved three times. These retrieved posters can be consulted

Figure 8. Initial state of one proposal for “The Prestige”

at: https://github.com/RmrezG/Five-cases-retrieved-for-The-Prestige.

Afterwards, the heuristics portion of the system established the acceptable ranges for the proposals’ qualitative features such as *positivity* (from 58.95% to 75.03%) and *type of symmetry* (as diagonal).

Given the previously mentioned features and inputs, CACARO initialized the characters’ positions and sizes in one of the proposals as shown in Figure 8. CACARO then began evaluating the proposal, starting with the relative proportions amongst the characters and the poster until they were within the acceptable limits. Afterwards, the system started making adjustments to reach the required positivity level, reaching 60.13%. Then, given the type of symmetry, CACARO tried to compensate the weights of each side of the symmetry axis until it achieved a balance of around 47%.

Finally, CACARO checked for overlapping faces and corrected positions when needed and applied the corresponding filter. It is important to notice that the evaluation process is sequential and iterative: when achieving an acceptable level of positivity, for example, the proportions criteria could end up being out of range even if it hadn't been originally. This forced CACARO to reevaluate and adjust the proposal until an acceptable value for all the criteria was reached, giving us the final state of this proposal as shown in Figure 9.

The adjustments consisted of resizing the characters (for the proportions and positivity criteria) or repositioning them (for the positivity, balance, and face overlap criteria). The magnitude and/or direction of these adjustments were determined based on optimizing the poster’s composition together with a random component, which allowed CACARO to avoid stagnation and to propose a variety of possible and creative solutions.

The degree of quality of this poster, obtained by the metrics previously detailed, were 100% for the proportions criterion, 89.76% for positivity, 97.00% for balance and 98.23% for face detection, resulting in an overall degree of quality of 96.24%.

The second test problem was the redesign of the poster for “Raiders of the Lost Ark” (Paramount Pictures 1981). The most important input requirements were “Action” (as *Genre*) and the image and features (*Class* and *Role*) of the main character. Therefore, CACARO made a poster composition based on one single character whose size and position determined the poster’s quality. The proportions, positivity, balance, and face detection criterion obtained quality values of 100%, 92.83%, 90.26% and 100%, respectively. The final poster, shown in Figure 10, obtained a quality level of 95.77%.

The final problem had as goal the redesign of the poster for the movie “When Harry Met Sally” (Columbia Pictures 1989), where the input requirements included the two main characters and the “Romance” genre. The proportions criterion and the face detection had a 100% value of quality,

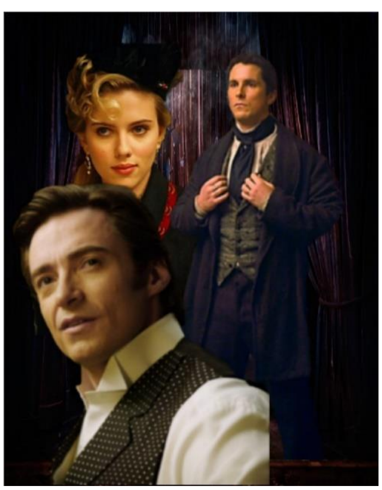


Figure 9. Final state of one proposal for “The Prestige”.

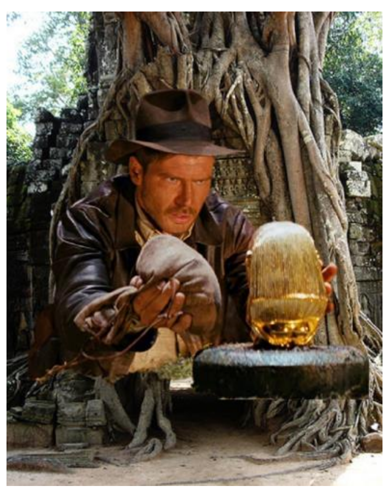


Figure 10. Final state of one proposal for “Raiders of the Lost Ark”.

Figure 11. Final state of one proposal for “When Harry Met Sally”.

while positivity and balance obtained 93.45% and 92.10%. The quality level for the final poster, shown in Figure 11, was 96.38%.

It is important to keep in mind that the whole process from generating an initial composition to getting the final result involves the encoding techniques previously described. Each poster proposal is described by its features (number of characters, genre, feeling, etc.), and each feature is represented by numbers and bit strings where each bit string follows the restrictions imposed by the one-hot or binary encoding representation. Those feature representations form the basis for generating different visual proposals.

Discussion, Conclusions and Future Work

In this paper we have described CACARO, a system that proposes movie poster compositions. The algorithms that CACARO follows in order to perform its task are centered around case-based reasoning, augmented with random forests (for case retrieval) and heuristic rules (for case adaptation). The combination of some of the random choices present in the system's process model and the limits imposed on them by its domain knowledge as embodied in the case memory provide the inspiration that the system uses in order to propose new poster compositions and are the source of its potential for creativity.

The current version of the system has some limitations which we plan to address in the future. Some of these are, for example, the ambiguity of some features. For instance, the feeling (mood) that the poster transmits and the role of a character in a movie may be subject to interpretation. The current implemented version of the system only has in its case base posters for movies in the horror, drama, romance, and action genres, which limits the range of values of the mood variable within the case memory, and this may affect the results.

On the one hand, a wider variety of genres may provide the system with the capability to produce poster compositions that are more "fine-tuned" to the particular specifications of a new problem. On the other hand, the more genres that are included, the greater the possibility for overlap/ambiguity between genres, and thus the greater the chances that the system's proposals might not be as good as possible. We plan to perform ablation experiments with a more complete case base and reduced subsets of the same case base (produced by extracting some genres) to see the effect this has on the types of solutions proposed by the system. There is also the possibility of talking with a panel of experts (composed of designers and filmmakers) to evaluate CACARO's results and try to tweak the poster generation process based on the experts' feedback. Another limitation is that our cases are structured in such a way that only up to three characters per poster may be included in the case description, and if there are less the values assigned to the absentees may disproportionately affect the random forest classifiers because we assign them a -10 value.

The size and scope of the case base (which contains 103 cases) might affect the system's predictions and accuracy. This results in some edge situations in which all the forests may return the same solution, and this doesn't leave room for changes to be made during the case adaptation phase of the system.

We intend to implement an evolutionary algorithm in order to make CACARO even more creative. The algorithm will start with a population consisting of the five cases resulting from the *random forests* plus some randomized individuals. Then, through many generations of mutation and crossover operations on the individuals we will generate more varied proposed compositions. To ensure the quality of each new generation our evaluation function will discard bad individuals and eventually will help converge on a result which will be the proposed poster. This will require that each individual in each generation be evaluated by its "phenotype", i.e., the poster itself. Thus, the system will require more computational power for image processing than the current version. We will also have to reformulate certain aspects of the evaluation function. Previous work (Gómez de Silva Garza and Maher 1999) has shown that evolutionary algorithms can provide good results when used for the adaptation of cases.

Acknowledgements

This work has been supported by Asociación Mexicana de Cultura, A.C.

References

Arnheim, R. 1974. *Art and Visual Perception: A Psychology of the Creative Eye (2nd ed.)*. Berkeley, California: University of California Press.

Breiman, L. 2001. Random forests. *Machine Learning* 45, 5–32.

Cohen, J., Cohen, P., West, S. G., and Aiken, L. S. 2013. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences.* London: Routledge.

Desolneux, A., Moisan, L., and Morel, J.-M. 2008. *From Gestalt Theory to Image Analysis: A Probabilistic Approach*. New York: Springer.

Gómez de Silva Garza, A., and Maher, M.L. 1999. An evolutionary approach to case adaptation. In Althoff, K.D., Bergmann, R., and Branting, L.K. (eds.), *Case-Based Reasoning Research and Applications: Proceedings of the Third International Conference on Case-Based Reasoning (ICCBR)*. Berlin: Springer. 162-172.

Guberman, S. 2015. On Gestalt Theory Principles. *Gestalt Theory* 37(1), 25-44.

Harris, D., and Harris, S. 2013. *Digital design and computer architecture*. California: Morgan Kaufmann.

Kobourov S.G., Mchedlidze T., and Vonessen L. 2015. Gestalt Principles in Graph Drawing. In: Di Giacomo E., Lubiw A., eds., *Graph Drawing and Network Visualization*. Cham, Switzerland: Springer. 558-560.

Kolodner, J. 1993. *Case-Based Reasoning*. California: Morgan Kaufmann.

Kubat, M. 2017. Decision trees. In *An Introduction to Machine Learning (2nd ed.)*. Cham, Switzerland: Springer. 113-135.
Leake, D. 1996. *Case-Based Reasoning: Experiences, Lessons, and Future Directions*. Cambridge, USA: AAAI Press/MIT Press.
Löw, N., Jürgen, H., and Blessing, M. 2019. Multiple retrieval case-based reasoning for incomplete datasets. In *Journal of Biomedical Informatics, Vol. 92.*
McManus, I.C., Stöver, K., and Kim, D. 2011. Arnheim's Gestalt Theory of Visual Balance: Examining the Compositional Structure of Art Photographs and Abstract Images. *i-Perception* 2, 615-647.
Pai, C., Pardawala, T. and Potdar, K: 2017. A comparative study of categorical variable encoding techniques for neural network classifiers. *International Journal of Computer Applications*, 175(4) 7–9.
Pearl, J. 1984. *Heuristics: intelligent search strategies for computer problem solving*. Reading, MA, USA: Addison-Wesley Pub. Co., Inc., p. 3.
Redmon, J., and Farhadi, A. 2018. YOLOv3: An incremental improvement. Tech Report available at arXiv:1804.02767.
Romanycia, M. H. J. and Pelletier, F. J. 1985. What is a heuristic? *Computational Intelligence*, 1(1), 47–58.
Suler, J. 2013. Negative space. In Zakia, R., *Perception and Imaging (4th ed.)*. Oxford: Focal Press (Elsevier). 18-23.
Shalev-Shwartz, S; and Ben-David, S. 2014. Understanding Machine Learning: From Theory to Algorithms. New York, NY: Cambridge University Press. 255-256.
Yang, Q., Wu, J. 2001. Enhancing the Effectiveness of Interactive Case-Based Reasoning with Clustering and Decision Forests. In *Applied Intelligence 14*. 49-64.
Yang S.; Luo, P.; Loy, C.C.; and Tang, X. 2016. WIDER FACE: A Face Detection Benchmark. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 5525-5533.