Gaze Transition Entropy

KRZYSZTOF KREJTZ, University of Social Sciences and Humanities, and National Information Processing Institute, Warsaw, Poland
ANDREW DUCHOWSKI, Clemson University, Clemson, South Carolina
TOMASZ SZMIDT, Institute of Fundamental Technological Research, Polish Academy of Sciences, Warsaw, Poland
IZABELA KREJTZ, University of Social Sciences and Humanities, Warsaw, Poland
FERNANDO GONZÁLEZ PERILLI, Center for Basic Research in Psychology, Universidad de la República Uruguay, and Department of Basic, Evolutionary and Educational Psychology, Universitat Autònoma de Barcelona, Spain
ANA PIRES, Center for Basic Research in Psychology, Universidad de la República Uruguay, and Department of Basic, Evolutionary, and Educational Psychology, Universitat Autònoma de Barcelona, Spain
ANNA VILARO, Center for Research in Ambient Intelligence and Accessibility in Catalonia (CAIAC), Universitat Autònoma de Barcelona, Spain
NATALIA VILLALOBOS, Center for Basic Research in Psychology, Universidad de la República Uruguay

4

This article details a two-step method of quantifying eye movement transitions between areas of interest (AOIs). First, individuals' gaze switching patterns, represented by fixated AOI sequences, are modeled as Markov chains. Second, Shannon's entropy coefficient of the fit Markov model is computed to quantify the complexity of individual switching patterns. To determine the overall distribution of attention over AOIs, the entropy coefficient of individuals' stationary distribution of fixations is calculated. The novelty of the method is that it captures the variability of individual differences in eye movement characteristics, which are then summarized statistically. The method is demonstrated on gaze data collected from two studies, during free viewing of classical art paintings. Normalized Shannon's entropy, derived from individual transition matrices, is related to participants' individual differences as well as to either their aesthetic impression or recognition of artwork. Low transition and high stationary entropies suggest greater curiosity mixed with a higher subjective aesthetic affinity toward artwork, possibly indicative of visual scanning of the artwork in a more deliberate way. Meanwhile, both high transition and stationary entropies may be indicative of recognition of familiar artwork.

Categories and Subject Descriptors: J.4 [**Computer Applications**]: Social and Behavioural Sciences—*Psychology*

General Terms: Human Factors

Additional Key Words and Phrases: Eye movement transitions, entropy, Markov chain

Authors' addresses: K. Krejtz and I. Krejtz, Chodakowska 19/31, 03-815 Warsaw, Poland; emails: {kkrejtz, ikrejtz}@swps.edu.pl; A. Duchowski, 309 McAdams Hall, Clemson, SC 29634 USA; email: duchowski@clemson.edu; T. Szmidt, Adolfa Pawinskiego 5b, 02-106 Warszawa; email: tomasz.szmidt@gmail.com; F. González Perilli, A. Pires, and N. Villalobos, Tristán Narvaja 1674. 11200, Montevideo, Uruguay; emails: {fernando.gonzalez, apires}@psico.edu.uy, natvillalobosmanriquez@gmail.com; A. Vilaro, Edifici Q - Campus UAB, 08193 Cerdanyola del Vallés, Barcelona, Spain; email: annavilaro@gmail.com.

DOI: http://dx.doi.org/10.1145/2834121

ACM Reference Format:

Krzysztof Krejtz, Andrew Duchowski, Tomasz Szmidt, Izabela Krejtz, Fernando González Perilli, Ana Pires, Anna Vilaro, and Natalia Villalobos. 2015. Gaze transition entropy. ACM Trans. Appl. Percept. 13, 1, Article 4 (December 2015), 20 pages.
DOI: http://dx.doi.org/10.1145/2834121

1. INTRODUCTION

There is a pressing need for quantitative comparison of eye movement metrics [Duchowski et al. 2010]. There are two leading eye movement metrics for which quantitative similarity (or difference) methods have been developed. The first are scanpaths, represented by an ordered sequence of fixations, for which vector and string-based editing approaches have been developed to compute similarity [de Bruin et al. 2013; Jarodzka et al. 2010]. The second are heatmaps, represented (generally) by Gaussian mixture models (GMMs) indicating frequency (or probability) of fixation localization. Lacking sequential order information, heatmap-based metrics generally involve a measure of overlap between two GMMs, indicating similarity of fixated regions of an image, but not their order of visitation [Grindinger et al. 2010a, 2010b].

Transition matrices are a third type of sequential gaze pattern analysis but for which quantitative similarity measures are scarce. In this article, we present a method for comparing transition matrices and distribution of fixations between areas of interest (AOIs). The utility of quantitative comparisons between gaze pattern representations usually lies in the technique's ability to produce tests of statistical significance.

We extend the work of Krejtz et al. [2014] by addressing the following important omitted details:

(1) handling sparse matrices, e.g., when no fixation transition from a given source AOI is observed,
(2) normalizing entropy,
(3) computing the stationary distribution estimate via eigenvalue analysis, and
(4) considering the effects of content-independent AOI grid granularity.

2. BACKGROUND

Ellis and Stark [1986] introduced first-order (fixation) transition matrices, transforming them to conditional probability matrices for which they calculated conditional information H_c, $H_c = -\sum_{i=1}^{n} p_i \sum_{j=1}^{n} p_{ij} \log_2 p_{ij}$, $i \neq j$, where p_i is the simple probability of viewing the i^{th} AOI, p_{ij} is the conditional probability of viewing the j^{th} AOI given the previous viewing of the i^{th} AOI, and n is the number of AOIs. The value H_c, or *entropy*, provides a measure of statistical dependency in the spatial pattern of fixations represented by the transition matrix and may be used to compare one matrix with another. Ellis and Stark compared transition matrices of airline pilots viewing a cockpit display of traffic information (CDTI). The CDTI was fixed with eight AOIs. Weiss et al. [1989] note that in a transition matrix, a small H_c suggests dependencies between the fixation points, whereas a large H_c suggests a random scanning pattern. H_c is analogous to *transition entropy* $\hat{H}_t$ derived later in (2), which we use for similar comparative purposes as Ellis and Stark. However, the key difference between our approach is that we consider self-transitions. We also use a *content-independent* AOI grid, which allows us to vary the number of AOIs irrespective of the content being viewed. The benefit of the content-independent approach is that it allows estimation of transition matrices irrespective of the expected AOIs in the scene. All that is required is knowledge of the dimensions of the screen to establish different grid granularities. This makes the statistical analysis code portable among different experimental designs.

Ponsoda et al. [1995] used the probability vector and transition matrix to categorize saccade directions. Interestingly, their matrices were based on transitions between eight cardinal (compass) saccade directions (e.g., N, NE, SE, S) instead of from and to AOIs, as is more commonly done today. Although they used Z and χ^2 statistics to compare matrices, they did not consider the sequence of saccade directions to be a Markov chain, nor did they employ any Markov chain properties to obtain conclusions about saccade directions. In this article, we focus on saccadic transitions between AOIs instead of cardinal directions and model them as Markov chains. A Markov chain is a stochastic system undergoing transitions between states, in which the next state depends only on the present one and does not depend on the past. Such random processes have found many applications in social, technical, and life sciences.

Goldberg and Kotval [1999] discussed transition matrices between AOIs as one of several eye movement comparison metrics in usability evaluation, including number of saccades, saccade length and duration, saccadic amplitude, convex hull area, spatial density, number of fixations, fixation durations, and a fixation/saccade ratio. They noted the difference between content-dependent AOI assignment to specific screen objects and content-independent AOI assignment where the screen is simply divided into a uniform grid. To compare matrices, they suggested transition matrix density—the number of nonzero matrix cells divided by the total number of cells—as an indicator of matrix sparsity. In this work, we develop transition matrix entropy as a means of matrix comparison. Entropy estimation is a histogram-based method, which in effect takes into account the width of the histogram bins (AOI size), an aspect that density alone may miss if it only considers the total number of cells.

Liechty et al. [2003] used hidden Markov models (HMMs) to differentiate between local and global covert visual attention states from eye movement data. Instead of applying the Markov model to transition matrices as we do in this article, their use of HMMs was aimed at distinguishing fixations, akin to Velichkovsky et al. [2005], who suggested classifying fixations as ambient or focal.

Bednarik et al. [2005] constructed transition matrices to describe visual switching behavior among predefined regions but found no correlation between task performance and this high-level data. Fischer and Peinsipp-Byma [2007] combined transition matrices and descriptive statistics extracted from a uniform grid sampling of the stimulus to compare individuals' perception of the stimulus. However, neither of these efforts compared transition matrices.

Hwang et al. [2011] introduced transitional semantic guidance computation to evaluate fixation transitions. Instead of constructing a transition matrix from AOIs or a grid, they considered transitions from locations within a generated saliency map of a given scene. Semantic saliency of the next fixation target determined the guidance score of that gaze transition. This method can be considered as something of a hybrid between transition matrix construction and scanpath comparison where transition matrices are replaced by semantic maps. One conceivable disadvantage of this approach is comparison between scenes since saliency maps generated for each are likely to differ. Constructing content-independent transition matrices would overcome this problem.

Finally, the most recent example of work similar to ours is that of Vandeberg et al. [2013]. They analyze eye movement transitions with a multilevel Markov modeling approach. However, when comparing transition matrices, they compare matrix entries (transition probabilities) directly, whereas we use the entropy of individual transition matrices as a means of transition matrix comparison (advantages of our approach are made clear in the following section).

Our approach consists of two entropy measurements: transition entropy (calculated for individual subjects' transition matrices) and stationary entropy (calculated for individual subjects' stationary distributions), allowing for achieving straightforward measures of predictability (reflexiveness) in AOI transitions and overall distribution of eye movements over stimuli. Moreover, these two measures can be easily employed for comparisons via common statistical procedures.

Although the usefulness of entropy for analyzing eye movements has been previously demonstrated, the manner of usage of this measure differs significantly [Althoff 1998; Acartürk and Habel 2012; Shic et al. 2008; Kruizinga et al. 2006]. For instance, Acartürk and Habel [2012] examined visual comprehension of graphs in multimedia environments and calculated entropy for summary transition matrices where the cells' values represent the number of transitions between certain AOIs. This method, however, did not allow for discovering significant differences or other relations via statistical tests. Similarly, Shic et al. [2008] used entropy measures to find differences in the visual exploration of faces by autistic and normally developing children. They, however, did not distinguish between transition and stationary entropy estimates.

The present work describes the method for achieving the two measures of entropy and their properties, and it shows their usefulness when applied to a number of empirical examples.

3. MARKOV PROPERTY OF GAZE TRANSITIONS

A common practice while analyzing saccadic transitions between AOIs is to calculate transition matrices for each group of subjects under given experimental conditions. The value in row i and column j denotes the empirical probability that the next fixation would be to the j^{th} AOI if it were presently in the i^{th} AOI.

In this work, we apply a different approach based on transition matrices calculated for every subject, not groups of them. The main advantage is that we can analyze individual switching patterns between AOIs, which can vary significantly between subjects. A limitation lies in the number of AOIs defined: if there are too many, rows of the transition matrix can be equal to zero due to an insufficient amount of data. This problem is addressed in the implementation notes that follow.

Consider one subject, and let X_t be a random variable taking values in the set of AOIs $\mathcal{S} = \{1, \ldots, s\}$ (the state space) and denotes where this subject focused attention during the t^{th} fixation $(t = 0, \ldots, n)$. Analyses of transition matrices require ensuring that the next AOI depends only on the current one and that it is not affected by the past. Precisely, we need

$$\mathbb{P}(X_{t+1} = x_{t+1} | X_t = x_t, \ldots, X_0 = x_0) = \mathbb{P}(X_{t+1} = x_{t+1} | X_t = x_t),$$

where $x_t \in \mathcal{S}$. This condition is called the *Markov property*, and a stochastic process satisfying this is called a first-order *Markov chain* (or just a Markov chain). When the future state depends on both the present state and the one before, we can define a second-order Markov chain, and so on. The process describing transitions of the subject's visual attention that follows the Markov property is fully determined by the initial state $X_0 = x_0$ and transition matrix $\mathbf{P} = (p_{ij})_{s \times s}$, where $p_{ij} = \mathbb{P}(X_{t+1} = j | X_t = i)$, $t = 0, \ldots, n-1$ is the probability of changing the fixation's gaze position from the state i to j. We shall validate the quality of fitting a temporally homogeneous Markov chain (i.e., switching probabilities that do not depend on t) to the observed individual switching patterns.

To our knowledge, the Markov property has not yet been tested in modeling AOI sequences. Let us describe the test procedure presented by Besag and Mondal [2013]. Suppose that we are given the Markov chain with transition matrix $\mathbf{P}$ and initial state x_0. By the chain rule, the probability of obtaining the observed sequence $\mathbf{x} = (x_0, \ldots, x_n)$ with frequency counts n_{ij} of transitions from state i to j is equal to the product $\prod_{i,j \in \mathcal{S}} p_{ij}^{n_{ij}}$, which is referred to as the likelihood of the given sequence. The common log-likelihood ratio test statistic takes the form of

$$u = -2 \log \frac{\prod_{i,j,k} \left(\frac{n_{+jk}}{n_{+j+}} \right)^{n_{ijk}}}{\prod_{i,j,k} \left(\frac{n_{ijk}}{n_{ij+}} \right)^{n_{ijk}}}, \tag{1}$$

where n_{ijk} is the frequency count of the AOI triple (i, j, k) in the sequence $\mathbf{x}$ and $+$ denotes the summation over all AOIs. The numerator in (1) is the likelihood of the data when we assume the null hypothesis that the sequence follows a Markov chain of order 1 and the denominator is the likelihood of the alternative—a Markov chain of order 2. A more complex model fits the data better, so the test statistic is positive and its high value suggests rejecting the assumption of the Markov property.

Notice that expression (1) can be transformed to the form $u = 2\sum_{i,j,k\in\mathcal{S}} O_{ijk} \log O_{ijk}/E_{ijk}$, where $O_{ijk} = n_{ijk}$ is the observed frequency and $E_{ijk} = n_{ij+}(n_{+jk}/n_{+j+})$ is the expected frequency of a triple (i, j, k) under the null hypothesis assumption. The test statistic is proportional to the Kullback-Leibler divergence [Kullback and Leibler 1951] of the observed distribution of triples from the theoretical when the Markov property holds.

When the null hypothesis is valid, the test statistic has an asymptotic distribution χ^2 with $s(s-1)^2$ degrees of freedom. AOI sequences are often short, causing the distribution of the test statistic to be far from the asymptotic limit. Besag and Mondal [2013] proposed an exact test, based on random sampling from the space of all sequences starting from the same state x_0 with transition counts n_{ij} to be the same as the observed sequence $\mathbf{x}$. An example of how to test the Markov property is given in the following applied example.

4. ENTROPY OF AOI SEQUENCE

Given a sequence $\mathbf{x} = (x_0, \ldots, x_n)$ modeled as a Markov chain describing the subject's AOI switching pattern, with constant transition probabilities p_{ij} and stationary probabilities π_i, where $i, j \in \mathcal{S}$. Estimation of p_{ij} and π_i is discussed in the following implementation notes. The complexity of the process can then be measured by Shannon's entropy [Ekroot and Cover 1993; Ciuperca and Girardin 2005]:

$$\hat{H}_t = -\sum_{i\in\mathcal{S}} \pi_i \sum_{j\in\mathcal{S}} p_{ij} \log_2 p_{ij}. \tag{2}$$

Maximum entropy equal to $\log_2 s$ is reached when the distribution of transitions is uniform for each AOI (which also means that subsequent transitions are independent of each other). The minimal entropy of 0 describes a fully deterministic Markov chain—that is, one in which all transition probabilities are either 0 or 1. Thus, the higher the entropy, the more randomness there is in a subject's transitions, the more complex the sequence of AOIs, and the more exploratory the character of visual attention.

Stated another way, entropy refers to the "expected surprise" of a given gaze transition. Minimum entropy of 0 suggests no expected surprise, meaning that a gaze transition is always expected to the same j^{th} AOI. Maximum entropy, on the other hand, suggests maximum surprise, since transition from source AOI to any destination AOI is equally likely, and hence whichever occurs results in maximum expected surprise. More formally, the term $-p_{ij} \log_2 p_{ij}$ in (2) is the transition's contribution to system entropy, modeled by its probability multiplied by its *surprisal* [Hume and Mailhot 2013].

Generally speaking, the entropy rate of a stochastic process $\mathcal{X} = (X_1, X_2, \ldots)$ is defined by using the entropy of the variables joint distribution $(X_1, \ldots, X_n)$, $H(\mathcal{X}) = \lim_{n\to\infty} \frac{1}{n} H(X_1, \ldots, X_n)$, when the limit exists. If the process is the stationary Markov chain then its entropy rate is equal to the transition entropy (2) [Cover and Thomas 2006], $H(\mathcal{X}) = \hat{H}_t$. We are also interested in the entropy of the stationary distribution itself:

$$\hat{H}_s = -\sum_{i\in\mathcal{S}} \pi_i \log_2 \pi_i. \tag{3}$$

A higher value of stationary entropy means that the subject distributes their visual attention more equally among AOIs. A lower value is obtained when fixations tend to be concentrated on certain AOIs, possibly because they appear to be more interesting to the viewer.

One can expect that the long-term distribution of fixations over AOIs should be more uniform than the distribution of transitions (i.e., stationary entropy is greater than transition entropy). That $\hat{H}_s \geq \hat{H}_t$ is a well-known property [Cover and Thomas 2006; Feixas et al. 1999]. If we treat a Markov chain as an information channel with input X, output Y, and conditional probabilities p_{ij}, where both X and Y have stationary distribution $\boldsymbol{\pi}$, then the entropy rate (information content) of the Markov chain $\hat{H}_t$ is equal to the conditional entropy $H(Y|X)$. On the other hand, conditioning reduces uncertainty, so $H(Y|X) \leq H(Y) = \hat{H}_s$.

5. IMPLEMENTATION

Transition matrix $\mathbf{P}$ can be computed in R [R Development Core Team 2011] for each of the AOIs defined atop the stimulus image. Matrix elements p_{ij} are set to the number of transitions from i^{th} source AOI to j^{th} destination AOI for each participant and then the matrix is normalized relative to each source AOI (i.e., per row),

$$p_{ij} = \frac{n_{ij}}{\sum_j n_{ij}}, \quad i, j \in \mathcal{S}, \tag{4}$$

such that p_{ij} represents the estimated probability of transitioning from i^{th} AOI to any j^{th} AOI given i^{th} AOI as the starting point.

5.1 Entropy Computation and Normalization

In practice, it is possible that no transitions from the i^{th} AOI are observed. This leads to a zero matrix row sum in (4) and division by zero. When this occurs, we handle this happenstance by setting each of the row entries to their uniform transition distribution, namely $p_{ij} = 1/s$, thereby modeling an equally likely probability of transitioning to any other AOI given this i^{th} source AOI. An added benefit of this decision is that it leads to the construction of a transition matrix $\mathbf{P}$ that is regular (specifically a right stochastic matrix), with all entries positive and nonzero, facilitating stationary entropy calculation via Eigen analysis (see later discussion). Note that setting each of the p_{ij} entries to $1/s$ would lead to a uniform matrix with maximum entropy equal to $\log_2 s$ bits per transition. Indeed, maximum entropy is used to normalize the empirical entropy obtained from each transition matrix. Empirical entropy is computed by the maximum likelihood method as implemented by R's publicly available `entropy` module [Hausser and Strimmer 2009].

5.2 Statistical Comparison

To facilitate statistical comparison of mean entropies per experimental condition, $\hat{H}_t$ is computed per individual participant and per condition and normalized, which yields $H_t = \hat{H}_t / \log_2 s$. This results in a table of entropies (each entropy computed from an individual's transition matrix) for each of experimental conditions and each of the participants. ANOVA is then used to test for differences in mean (normalized) entropy per condition.

5.3 Stationary Entropy Implementation via Eigen Analysis

Stationary entropy $\hat{H}_s$ and stationary distribution $\boldsymbol{\pi}$ are estimated through eigenvalue analysis. Transition matrix $\mathbf{P}$ represents the likelihood of the state of the system's next period. The likelihood of the state in two periods is given by $\mathbf{P}_2 = \mathbf{PP}$. Similarly, the likelihood of the state in the long run is given by $\mathbf{P}_n = \mathbf{P}^n$. If the transition matrix $\mathbf{P}$ is regular, then over a sufficient number of periods n, the system approaches a fixed state and transition probabilities will converge to $\boldsymbol{\pi}$—that is, $\mathbf{xP}^n \rightarrow \boldsymbol{\pi}$, where vector $\mathbf{x}$ denotes an arbitrary probability distribution. Here, the regularity means that some power of matrix $\mathbf{P}$ has all of the components positive, which means that any AOI can be reached from any AOI after some transitions. This assumption has been verified in the subsequent analyses. The steady-state (stationary) probability vector $\boldsymbol{\pi}$ can be solved for directly if the following holds: $\boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}$.

Given transition matrix $\mathbf{P}$, R's `eigen` module returns both of its eigenvalues and eigenvectors. The first eigenvector corresponds to eigenvalue equal to 1, and its normalization gives the stationary distribution, with normalized stationary entropy computed as $H_s = \hat{H}_s / \log_2 s$, where $\hat{H}_s$ is given by (3).

5.4 Example Applications of Transition Entropy Analysis

The utility of our entropy-based metrics for analysis of eye movement transitions is demonstrated by two practical examples of viewing artwork. In the first example, artwork from three periods is viewed by observers, and their gaze transitions are evaluated in relation to their personal curiosity and interest in the artwork. In the second, we compare viewers' gaze transitions over a stylized rendering of *La Jaconde* in relation to their self-reported recognition of the painting's representation. In both examples, we consider the effects of (content-independent) AOI grid granularity. The first example splits the image vertically, and the second tessellates the image into rectangular (and uniform) grid cells.

6. STUDY 1: VIEWING ARTWORK IN THREE CLASSICAL STYLES

The first study compared gaze transitions from viewing paintings from three classical periods: Impressionist, Renaissance, and Bauhaus. Specifically, three paintings (digital reproductions of oil paintings) were used as stimulus: "Jour de pluie a Paris" by Gustave Caillebotte (Impressionism), "The Tempest" by Giorgione (Renaissance), and "Composition VII" by Vassily Kandinsky (Bauhaus). The choice of the stimuli was mainly motivated by the need for wide variability and discrepancy in artistic style. The chosen paintings were treated as exemplary representatives of the different styles. Moreover, the paintings represent different levels of abstraction: from the very figurative by Caillebotte to the very abstract by Kandinsky. We analyze data from a study conducted by Biele et al. [2013], wherein gaze data was analyzed in terms of ambient or focal fixations, as per Velichkovsky et al. [2005]. Our entropy-based transition analysis is newly applied to their data.

Experimental design and procedure. The free viewing experiment consisted of a presentation of the paintings series on the computer screen. An example of the artwork, with two and three AOIs defined for the analysis (marked with green lines) is shown in Figure 1.[1] AOIs on the remaining stimuli were marked in a similar way.

The experimental procedure consisted of two steps:

(1) First, all participants filled out the Curiosity and Exploration Inventory (CEI-II) [Kashdan et al. 2009] in paper and pencil form. The CEI-II is a recognized self-report test to assess the recognition, pursuit, and integration of challenging, novel experiences. It consists of 10 items to which answers are given on 5-point Likert-type scale.
(2) Second, participants were asked to view art paintings displayed in random order on the computer screen. Their task was "to examine artwork as they would look at them in an art album." Each stimulus was presented for 30 seconds in full-screen mode. Participants' eye movements were

[1] The picture of the artwork was retrieved from http://en.wikipedia.org/wiki/File:Giorgione_019.jpg. Source/photographer: Florian Heine *Das erste Mal. Wie Neues in die Kunst kam*, München: Bucher Verlag 2007, ISBN 978-3-7658-1511-9.

(a) Renaissance–2 AOIs *(b) Renaissance–3 AOIs*

Fig. 1. An example of Renaissance stimulus with two and three AOIs marked as used in the experiment.

recorded during this stage of the experiment. After each artwork presentation, they evaluated its attractiveness on a 5-point Likert-type scale.

Participants. Forty-nine university students (31 male and 18 female, aged $M = 24.47$, $SD = 6.53$) took part in the experiment after signing a consent form. Data from nine subjects were excluded from the analysis due to eye tracking system calibration problems resulting in a low tracking ratio or due to an insufficient amount of fixations.

Apparatus. All stimuli were presented on a computer monitor (1680×1050 resolution; 22-inch LCD, 60Hz refresh rate) connected to a standard PC. Eye movements were recorded at 120Hz with an SMI eye tracking system. SMI's BeGaze software was used for fixation and saccade detection with a dispersion-based event detection algorithm. The dispersion was 100px, with minimum fixation duration set to 80ms. When analyzing visual exploration patterns, it is frequently accepted to use fixations shorter than 100ms [Velichkovsky et al. 2005].

6.1 Results

Analysis of the eye movement data focuses on comparison of subjects' transition and stationary entropies in relation to their self-reported characterization of curiosity and their evaluation of the attractiveness of the freely viewed artwork.

6.1.1 *Results of Markov Chain Fitting.* For each stimulus and number of AOIs, we checked the resultant AOI sequences as to whether modeling with a Markov chain was valid. Results of the exact and asymptotic test proposed by Besag and Mondal [2013], described in Section 3, show that the assumed model fits the data reasonably well. Only a few cases occurred for which the hypothesis of the Markov property should be rejected (at a significance level of 0.05). This was confirmed by the test for multiple sequences—for all three stimuli the probability was close to 1, giving no reason to doubt that AOI sequences are independent and come from Markov chains with individual transition matrices.

Before computing the stationary entropy, we also tested the number of iterations required for the probabilities in the transition matrices to converge (to a stationary distribution). This was tested for each participant, stimuli, and grid granularity (two or three AOIs) by iterative multiplication of

Table I. Descriptive Statistics of Number of Iterations Needed to Converge Transition Probabilities into a Stationary Distribution

Artwork Style	Two AOIs	Three AOIs
Bauhaus	M = 15.6, SD = 9.95	M = 19.4, SD = 10.58
Impressionism	M = 17.98, SD = 6.84	M = 21.15, SD = 9.92
Renaissance	M = 10.1, SD = 3.25	M = 17.88, SD = 7.19

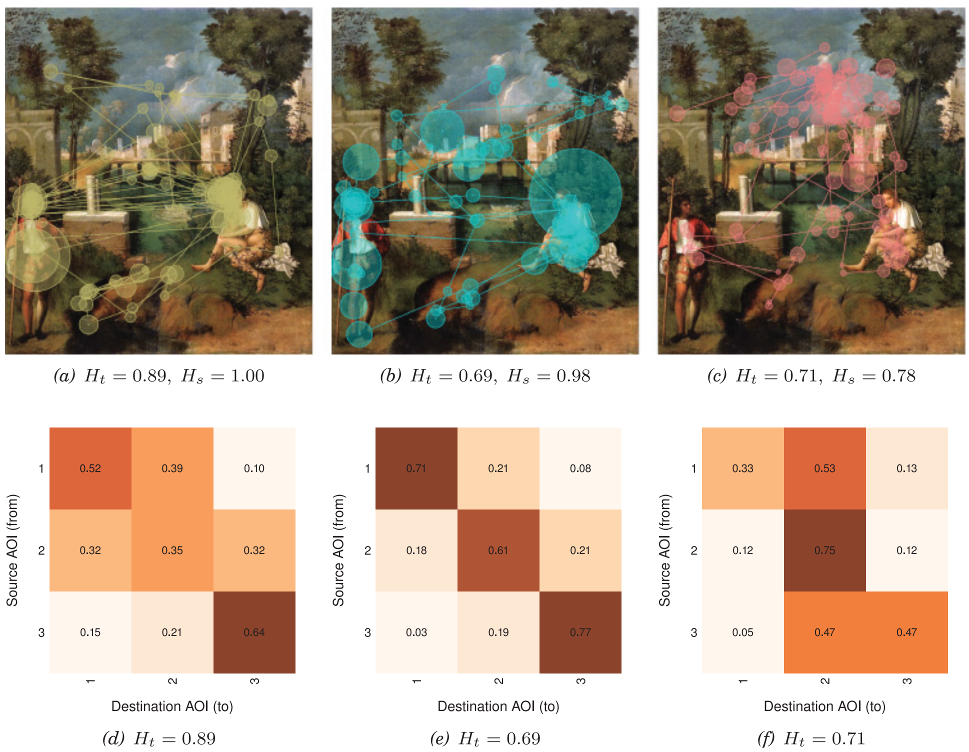
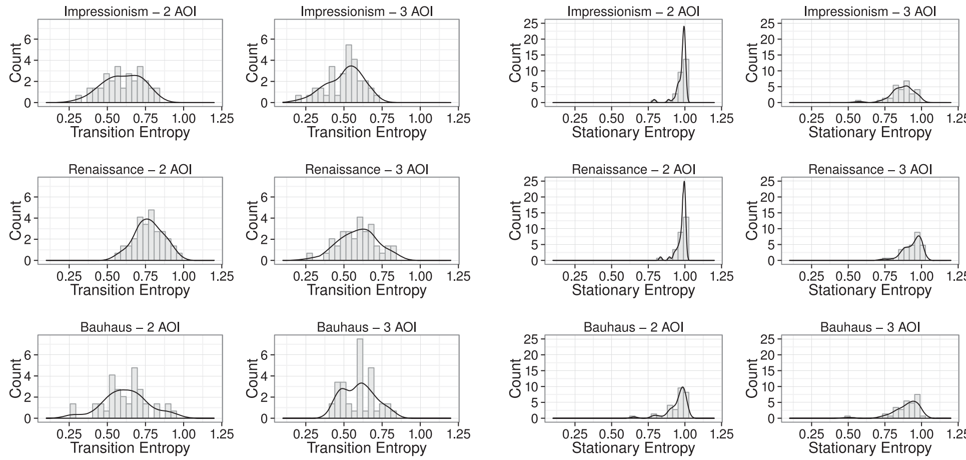


Fig. 2. Scanpaths (a) through (c) and transition matrices (d) through (f) with low and high entropy values: transition matrices are 3 × 3 reflecting transitions from each of the three source AOIs to the three destination AOIs.

transition matrix ($\mathbf{P}^n$ as $n \to \infty$). Iterative multiplication was stopped when the k^{th} iteration reached convergence according to criterion $|\boldsymbol{\pi}_i - p^k_{0i}| < 0.001,\ i \in \mathcal{S}$ or when the number of iterations reached 10,000 (no convergence). In all cases, transition probabilities converged into a stationary distribution. Descriptive statistics show that overall, the transition probabilities converge into stationary distribution in approximately 17 iterations, on average, with SD = 8.98. Detailed descriptive statistics for each test are presented in Table I.

6.1.2 *Transition and Stationary Entropies.* Scanpaths of three subjects with strongly differing entropy values are shown in Figure 2. The first two images present AOI sequences with a similar stationary entropy H_s and different value of H_t. In Figure 2(a), the subject frequently switched between male and



(a) Transition entropies *(b) Stationary entropies*

Fig. 3. Distributions of entropies for different stimuli and number of AOIs.

Table II. Results of Shapiro-Wilk Test of Normality Distribution of Transition and Stationary Entropy

	Transition Entropy H_t		Stationary Entropy H_s	
Artwork Style	Two AOIs	Three AOIs	Two AOIs	Three AOIs
Bauhaus	$W = 0.98$, $p > 0.1$	$W = 0.96$, $p > 0.1$	$W = 0.78$, $p < 0.001$	$W = 0.79$, $p < 0.001$
Impressionism	$W = 0.98$, $p > 0.1$	$W = 0.97$, $p > 0.1$	$W = 0.57$, $p < 0.001$	$W = 0.83$, $p < 0.001$
Renaissance	$W = 0.99$, $p > 0.1$	$W = 0.98$, $p > 0.1$	$W = 0.74$, $p < 0.001$	$W = 0.79$, $p < 0.001$

female figures, resulting in high transition entropy. The low value of H_t indicates more careful viewing of AOIs, which is reflected in longer subsequences of fixations in the same area (Figure 2(b)).

The last two images present scanpaths with similar switching complexity but with highly differing stationary AOI distributions. In Figure 2(c), the subject was mostly interested in the sky rather than the woman nursing her child and paid little attention to the man. In Figure 2(b), the subject's visual attention was split more evenly between the AOIs, yielding a higher value of H_s.

Transition matrices of three selected scanpaths are depicted in Figure 2(d) through (f). More uniform transition probabilities in subsequent rows result in a higher value of H_t.

6.1.3 *Analysis of Observed Entropy.* Figure 3(a) and (b) show the distribution of H_t and H_s for all subjects viewing the Renaissance artwork. All distributions of H_t appear Gaussian in shape as tested statistically with a series of Shapiro-Wilk tests. In contrast, H_s distributions appeared to deviate significantly from normality. The detailed results of these tests are presented in Table II.

The results of the Shapiro-Wilk tests for H_t show that its distribution meets one of the most important assumptions for usage of parametric statistical procedures (e.g., ANOVA; see Figure 3(a)). For H_s, all distributions deviated from a normal distribution (see Figure 3(b)). For this reason, we recommend considering nonparametric statistical tests or procedures for dealing with stationary entropy data as a dependent variable.

Recall that in the case of a uniform distribution over the subset $\{1, \ldots, s\}$, the entropy is equal to $\log s$. This leads to the prediction that significant differences in mean values of the entropy should

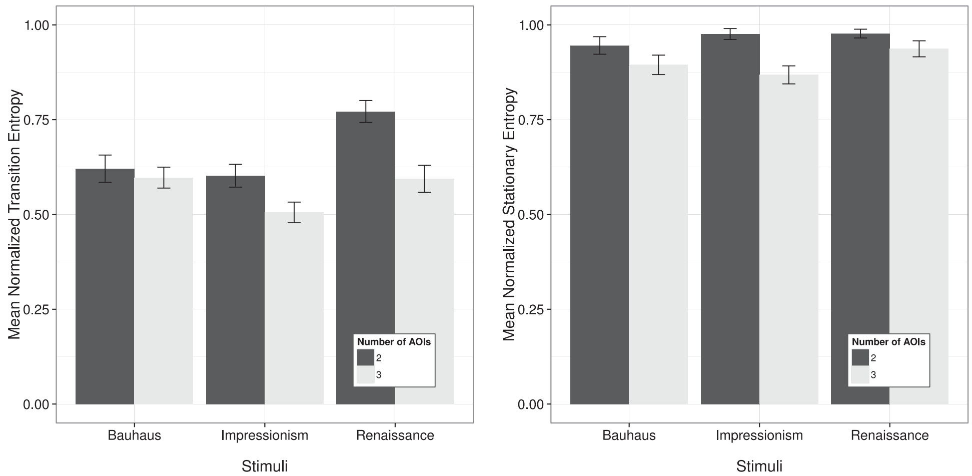


Fig. 4. Stationary and transition entropies versus number of AOIs and stimulus. Error bars represent ±95% confidence intervals.

be observed between stimuli with two and three AOIs. Moreover, we suspected significant differences between the stimuli, reflecting various art styles and content. To test these hypotheses, a within-subjects 2×3 ANOVA was performed. In this analysis, the number of AOIs (two vs. three) and stimuli (Impressionism vs. Renaissance vs. Bauhaus) were tested as within-subject factors and the transition entropy values (H_t) as the dependent variable.

In line with the hypothesis, analysis revealed a statistically significant main effect of number of AOIs, ($F(1, 41) = 100.94$, $p < 0.001$, $\eta_p^2 = 0.155$). Mean normalized transition entropy (H_t) for two AOIs was significantly higher (M = 0.667, SE = 0.01) than the mean transition entropy value for three AOIs (M = 0.566, SE = 0.01).

Results showed a significant main effect of stimuli ($F(2, 82) = 20.36$, $p < 0.001$, $\eta_p^2 = 0.148$). The mean transition entropy values (H_t) were the highest for the Renaissance stimulus (M = 0.679, SE = 0.02), lowest for the Impressionist (M = 0.559, SE = 0.02), and in-between for the Bauhaus (M = 0.610, SE = 0.01) stimuli. Pairwise comparisons with Bonferroni correction showed that all differences were statistically significant ($p < 0.01$).

The analysis also showed the interaction effect to be significant, $F(2, 82) = 24.45$, $p < 0.001$, $\eta_p^2 = 0.072$ (Figure 4). Pairwise comparisons with Bonferroni correction showed that for two AOIs, normalized transition entropy (H_t) from the Renaissance painting (M = 0.772, SE = 0.01) differs significantly (p < 0.001) from entropy from the Bauhaus (M = 0.622, SE = 0.02) and Impressionist paintings (M = 0.607, SE = 0.01).

For three AOIs, the entropy from the Impressionist painting (M = 0.512, SE = 0.01) differs significantly (p < 0.01) from the Renaissance (M = 0.588, SE = 0.02) and the Bauhaus paintings (M = 0.598, SE = 0.01).

The next set of post hoc tests revealed that the transition entropy for two AOIs is significantly higher (p < 0.001) than for three AOIs while looking at the Renaissance and the Impressionist artworks. There was no difference for the Bauhaus painting in transition entropies with two and three AOIs. These differences may be due to the nature of the paintings. Whereas Renaissance and Impressionist paintings were examples of figurative art with depiction of people and objects, the Bauhaus artwork was purely abstract.

Similar ANOVA was performed for normalized stationary entropy (H_s) as the dependent variable. Again, the number of AOIs and stimuli were treated as within-subject factors. The analysis revealed a significant main effect of the number of AOIs ($F(1, 41) = 77.22$, $p < 0.001$, $\eta_p^2 = 0.199$). The mean stationary entropy for two AOIs was significantly higher (M = 0.966, SE = 0.01) than for three AOIs (M = 0.900, SE = 0.01).

The main effect of stimuli reached significance level ($F(2, 82) = 5.39$, $p < 0.001$, $\eta_p^2 = 0.053$). The following pairwise comparisons showed that stationary entropy (H_s) on Renaissance artwork was significantly (p < 0.01) higher (M = 0.955, SE = 0.01) than on Bauhaus (M = 0.920, SE = 0.01) or Impressionist paintings (M = 0.924, SE = 0.01).

Perhaps more interestingly, the analysis revealed a significant interaction effect of number of AOIs and visual stimuli ($F(2, 82) = 8.55$, $p < 0.001$, $\eta_p^2 = 0.044$). The pairwise comparison with Bonferroni correction showed that for all stimuli treated separately, stationary entropies of two AOIs were significantly higher (p < 0.01) than for three AOIs (see Figure 4). Comparison of entropies calculated for two AOIs showed that the Bauhaus painting produced a significantly lower mean value of stationary entropy (M = 0.945, SE = 0.01) than the Renaissance (M = 0.975, SE = 0.01) and Impressionist artworks (M = 0.977, SE = 0.01). Visual attention to the figurative paintings (from the Renaissance and Impressionist periods) of viewers was more equally distributed among both AOIs than for the abstract Bauhaus art piece. Lower stationary entropy values (H_s) indicate that participants tend to focus more on specific parts of the Bauhaus painting when viewing it.

For three AOIs, results of pairwise comparisons showed that stationary entropy was significantly higher on the Renaissance painting (M = 0.934, SE = 0.01) compared with the Impressionist (M = 0.873, SE = 0.01) and Bauhaus stimuli (M = 0.895, SE = 0.01) (see Figure 4). These difference may again be explained by the content of the paintings. For the Renaissance painting, all AOIs contained distinct and clear figures (a boy, a woman, and a city), whereas all AOIs of the more modern paintings (Bauhaus and Impressionist) contained less unique content in visual terms.

In the context of eye movement analysis, what is perhaps most relevant is the use of transition matrix entropy to allow comparison of deployment of visual attention to expose differences in visual processing of potentially disparate stimuli. The preceding results show that the granularity of AOIs influences results—a point we examine in the following second study.

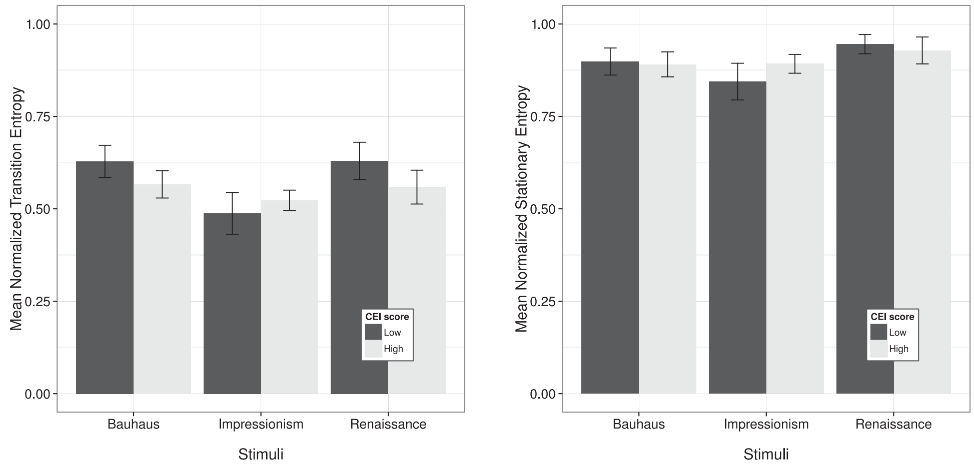
6.1.4 *Curiosity and Subjective Attractiveness Captured by Transition Entropy.*

One of the advantages of the proposed approach is the possibility of a statistical description of the dependence of eye movement transitions on individual differences. To demonstrate this, we tested two hypotheses assuming that eye movement transitions while looking on the art piece depend on the viewer's curiosity and her subjective evaluation of the piece.

In the first hypothesis, we expected specifically that highly curious participants would exhibit more focused eye movement patterns (i.e., concentrating longer on the same AOI, suggesting low transition entropy values), whereas less curious viewers would move their visual attention in a more chaotic way (switching frequently between different AOIs suggests higher transition entropy).

To test the curiosity hypothesis a 2×3 mixed-design ANOVA was conducted with H_t as the dependent variable. The curiosity score (CEI-II) was tested as a between-subject factor and stimuli as a within-subject factor. The analysis was run for three-AOI entropies.

The analysis revealed significant interaction between curiosity and stimuli, $F(2, 76) = 3.84$, $p < 0.05$, $\eta_p^2 = 0.045$ (Figure 5(a)). Although the interaction is significant its effect size is relatively low, which is also reflected in pairwise comparisons with Bonferroni correction, revealing only marginally significant differences. Participants with high curiosity scores had lower H_t values than those with low curiosity while looking at the Bauhaus ($p = 0.06$) and Renaissance ($p = 0.07$) paintings. For the



(a) Differences in H_t depending on curiosity and stimuli *(b) Differences in H_s depending on curiosity and stimuli*

Fig. 5. Differences in H_t and H_s depending on curiosity and painting style. Error bars represent ±95% confidence intervals. Analyses were done on data from three AOIs.

Impressionist painting, the difference between participants with high and low CEI scores was not significant.

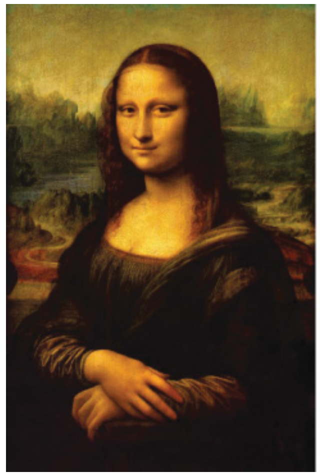
The analysis also showed the main effect of stimuli, $F(2, 76) = 12.1$, $p < 0.001$, $\eta^2 = 0.130$. The lowest transition entropy score was obtained for the Impressionism painting (M = 0.51, SE = 0.01) compared to the Bauhaus (M = 0.6, SE = 0.02) and Renaissance paintings (M = 0.59, SE = 0.02). The main effect of curiosity was not statistically significant $F(1, 38) = 1.57$, $p > 0.1$.

The preceding results support our prediction that highly curious participants would produce lower transition entropy while viewing the artwork. This can mean that their pattern of visual scanning is more locally oriented. They tend to carefully scan the first AOI, devoting many fixations to it, before moving on to the next. Thus, it is important to verify whether curious participants distribute their attention over the whole painting or just focus on a selected area. The latter would result in low H_s, which describes the long-term distribution of gaze transitions among all AOIs.

Considering this question, we conducted an ANOVA with mixed design 2×3 but with H_s as the dependent variable. Curiosity and stimuli were the independent variables. Results showed that curiosity does not significantly modify stationary entropy while looking at the different art styles, $F(2, 76) = 2.15$, $p > 0.1$ (see Figure 5(b)).

Analysis revealed a significant main effect of stimuli, $F(2, 76) = 8.30$, $p < 0.001$, $\eta_p^2 = 0.119$. Post hoc analyses showed that stationary entropy was significantly greater when looking at the Renaissance painting (M = 0.94, SE = 0.01) compared to the Bauhaus (M = 0.89, SE = 0.01) and Impressionist paintings (M = 0.89, SE = 0.01). The result means that visual attention was more equally distributed over all AOIs while viewing the Renaissance painting. The main effect of curiosity was not statistically significant.

The differences between subjects' eye movement transitions may also coincide with their evaluation of the artwork. Those who prefer an image may scan it with their eyes more carefully (i.e., in a more predictable way), resulting in lower H_t. A greater appreciation of the artwork may also lead to a more careful inspection of preferred regions of the painting, suggesting a relation between H_s and the subjective attractiveness of the art. Both hypotheses were tested with Pearson's correlations



(a) La Jaconde *(b)* La Jaconde *stylized*

Fig. 6. Painting and its point-like representation.

between subjective attractiveness and H_t and H_s. As expected, H_s significantly correlated with subjective attractiveness, $r = -0.37$, $t(34) = 2.28$, $p < 0.05$. However, contrary to our prediction, the correlation of H_t with subjective attractiveness was not statistically significant, $r = -0.14$, $t(34) = 0.82$, $p > 0.1$.

7. STUDY 2: VIEWING *LA JACONDE*

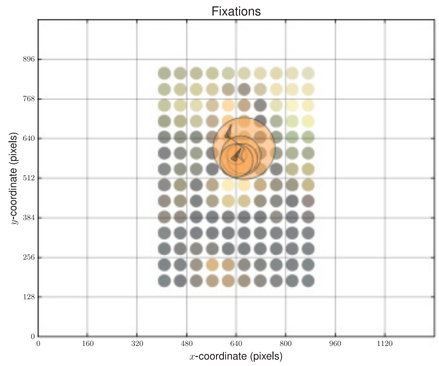
The second study was designed to evaluate the effect of different regular grid dimensions. The first study used 1×2 and 1×3 grids; here, we examine six different uniform grids and their effect on entropy, comparing gaze transitions when viewing a stylized image of *La Jaconde* (Figure 6).

Experimental design and procedure. The hypothesis behind this experiment posited that viewing differences would exist between those who recognized the stylized rendering of the painting and those who did not. The experiment proceeded with participants viewing the stylized representation of *La Jaconde* for 7 seconds ("1st view"). They were asked to press the space bar as soon as they recognized the painting. Those participants who said they recognized the painting were asked to indicate verbally what they thought they had recognized. They then viewed the digital representation of the painting in its original form, followed once again by the stylized version ("2nd view"). Differences in visual behavior between those who recognized the abstract version of the painting and those who did not would thus be expected in gaze data collected over the "1st view" of the stylized picture.

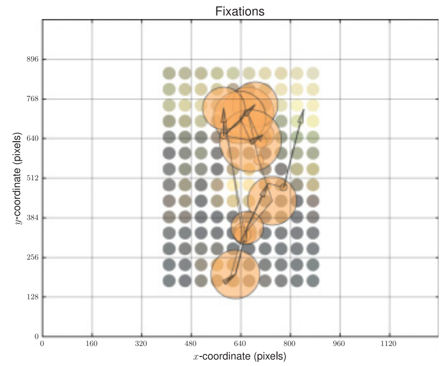
Participants. Thirty-one participants took part in the study. Of those, 17 self-reported that they did not recognize the abstract version of the painting, whereas 14 claimed they did.

Apparatus. Gaze data was recorded by a Tobii T60 (60Hz) eye tracker while participants viewed a 17-inch monitor (set to 1280×1024 resolution) at a viewing distance of about 20 inches.

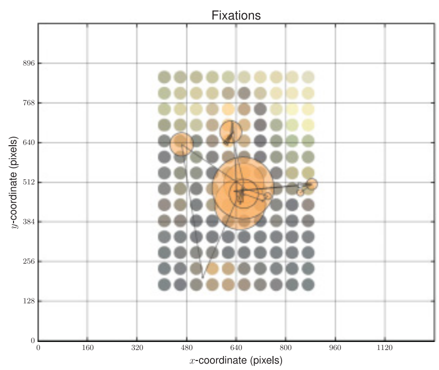
Raw data collected from the tracker was smoothed with a Butterworth filter set to 60Hz and 5.65Hz sampling and cutoff frequencies, respectively. Smoothed data was then convolved with a Savitzky-Golay differential filter acting as a velocity-based saccade detector (7-tap filter width with velocity threshold set to $\pm 5^\circ$/s) (see Duchowski et al. [2014]).



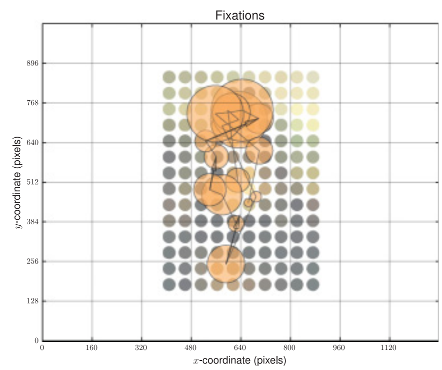
(a) *Viewer no. 28, who recognized* La Jaconde *1st view*; $H_t = 0.86$, $H_s = 0.96$



(b) *Viewer no. 18, who did not recognize* La Jaconde *1st view*; $H_t = 0.64$, $H_s = 0.89$



(c) *Viewer no. 28, who did recognize* La Jaconde *2nd view*; $H_t = 0.68$, $H_s = 0.91$



(d) *Viewer no. 18, who did not recognize* La Jaconde *2nd view*; $H_t = 0.58$, $H_s = 0.86$

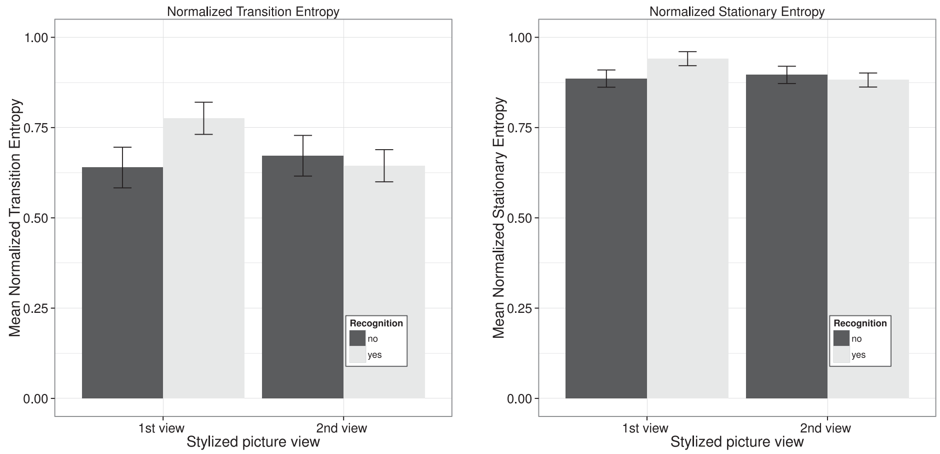
Fig. 7. Examples of scanpaths on stylized paintings ("1st and 2nd view") from two viewers: one who recognized and one who did not recognize *La Jaconde*.

7.1 Results

For analyses, we used an 8×8 AOI grid limited to the region of the stimuli with white spaces cut off around the stylized *La Jaconde* image (on the x-axis between 320 and 960 pixels, and on the y-axis between 128 and 896 pixels—c.f. Figure 7).

We first checked whether each of the AOI individual sequences can be modeled as a Markov chain given the AOI state space actually observed by the individual (i.e., only the few AOIs from the grid fixated by the individual). Similarly to the previous analysis, for all participants and experimental conditions, only a few cases did not meet the assumption of short-term process memory, rejecting the modeling assumption at a significance level of $p = 0.05$

Analysis of the eye movement data focuses on comparison of subjects' transition and stationary entropies, in relation to their self-reported identification of the stylized stimulus during the first and second viewings. We conducted two 2×2 analyses of variance, where the order of stylized picture viewing was used as a within-subject factor (1st vs. 2nd view) and recognition of *La Jaconde* as a between-subject factor (yes vs. no). For the first analysis, normalized transition entropy was used as the dependent variable and for the second, normalized stationary entropy.



(a) Normalized transition entropies vs. recognition and viewing order

(b) Normalized stationary entropies vs. recognition and viewing order

Fig. 8. Transition and stationary entropies versus viewing task or versus recognition. Error bars represent ±95% confidence intervals.

The normality of distributions of both normalized transition and stationary entropies were tested with a series of Shapiro-Wilk tests, separately for each stimulus (stylized *La Jaconde* picture). The tests showed that for each stimulus, the distributions did not significantly deviate from normality (in all cases $W \geq 0.95$, $p > 0.1$).

7.1.1 *Transition and Stationary Entropies.* Analysis of normalized transition entropy revealed a statistically significant effect of interaction between viewing order and recognition, $F(1, 28) = 8.09$, $p < 0.01$, $\eta^2 = 0.115$ (Figure 8(a)). Pairwise comparisons showed that during the first viewing of the stylized *La Jaconde* picture, participants who recognized it were characterized by a higher transition entropy of their eye movements ($H_t = 0.78$, $\mathrm{SE} = 0.02$) than those who did not ($H_t = 0.64$, $\mathrm{SE} = 0.02$). Additionally, participants who recognized the stylized *La Jaconde* during the first viewing exhibited significantly higher transition entropy ($H_t = 0.78$, $\mathrm{SE} = 0.02$) than during the second viewing ($H_t = 0.67$, $\mathrm{SE} = 0.02$).

Neither main effects of recognition, $F(1, 28) = 2.85$, $p > 0.05$, $\eta^2 = 0.045$, nor viewing order, $F(1, 28) = 1.50$, $p > 0.05$, $\eta^2 = 0.028$, were statistically significant.

Although the main effect of recognition was not significant, one possible interpretation of the preceding results is based on the uncertainty of the recognition process during the first viewing of the stylized picture. We may assume that recognition is preceded by more unpredictable eye movements (e.g., scanning the picture piecemeal), producing higher transition entropy during confirmation of recognition. This assumption posits that scanning of the picture prior to its recognition also produces higher stationary entropy (more balanced distribution over the AOIs).

ANOVA on stationary entropy as a dependent variable was performed to test the preceding assumption. Results revealed a statistically significant interaction effect, $F(1, 28) = 9.00$, $p < 0.01$, $\eta^2 = 0.147$ (see Figure 8(b)). Post hoc comparisons showed that for the first viewing of the stylized *La Jaconde* picture, participants who recognized it produced statistically higher stationary entropies ($H_s = 0.94$, $\mathrm{SE} = 0.01$) than those who did not recognize the picture ($H_s = 0.90$, $\mathrm{SE} = 0.01$). The difference between the first and second viewings of the stylized *La Jaconde* picture was significant for those who recognized it in the first viewing ($H_s = 0.90$, $\mathrm{SE} = 0.01$). This result suggests that the lower predictability of

Table III. Effect of Grid Size on Inferential Statistics

AOI Grid Size	Interaction Viewing Order × Recognition
2×2	$F(1, 28) = 3.46$, $p = 0.07$, $\eta^2 = 0.049$
4×4	$F(1, 28) = 0.86$, $p = 0.36$, $\eta^2 = 0.013$
8×8	$F(1, 28) = 8.09$, $p < 0.01$, $\eta^2 = 0.115$
10×8	$F(1, 28) = 0.86$, $p = 0.36$, $\eta^2 = 0.011$
16×16	$F(1, 28) = 0.02$, $p = 0.90$, $\eta^2 = 0.001$
20×16	$F(1, 28) = 0.88$, $p = 0.36$, $\eta^2 = 0.007$

viewers' gaze transitions atop the familiar image of *La Jaconde* carries on into the long term (steady' state).

Similarly to the analysis of transition entropy, in the analysis of stationary entropy, neither main effect of recognition, $F(1, 28) = 2.39$, $p > 0.05$, $\eta^2 = 0.038$, nor viewing order, $F(1, 28) = 2.42$, $p > 0.05$, $\eta^2 = 0.044$, was significant.

7.1.2 *Effect of Grid Granularity.* Grid granularity has a clear impact on the transition entropy statistics and therefore must be carefully considered prior to testing transition entropy for statistical significance and prior to drawing inferential conclusions from such results.

To illustrate the impact of grid granularity on transition entropy statistics, we repeated the analysis on five additional grid dimensions including the 8×8 grid: 2×2, 4×4, 8×8, 10×8, 16×16, and 20×16.

Focusing on normalized transition entropy, only the 8×8 grid showed a statistically significant interaction between viewing order and recognition, as reported earlier. The complete list of significance tests is given in Table III, with transition entropies shown graphically in Figure 9(a).

8. DISCUSSION AND CONCLUSIONS

We introduced a new method for the analysis of gaze transitions between different AOIs. Gaze transitions are accumulated from source to destination AOIs defined over the stimulus image, then normalized relative to each source AOI (i.e., per row of the transition matrix). Statistical analysis is derived from modeling gaze transitions as first-order Markov chain processes. This assumption is tested via the statistical test of Besag and Mondal [2013], used to confirm the validity of modeling transitions as short, first-order processes.

The complexity of both transition and stationary distributions between AOIs is expressed in terms of normalized Shannon's entropy. Higher transition entropy H_t denotes more randomness and more frequent switching between AOIs. Consequently, individual switching patterns between AOIs can be visualized and statistically compared. Furthermore, the stationary entropy of gaze transitions H_s depicts the extent to which individual visual attention is distributed among AOIs. Higher stationary entropy H_s means more uniform distribution of visual attention over selected AOIs. Each normalized entropy coefficient, H_t and H_s, ranges from 0 to 1, with 0 indicating highest predictability and 1 complete randomness.

Entropy-based analysis of gaze transitions was demonstrated on two empirical eye-tracking studies. The validity of modeling gaze transitions as short, first-order (one fixation memory) processes was verified independently by the test of Besag and Mondal [2013] on data from both studies. Transition and stationary entropy measures were then used to describe differences in individuals' visual attention switching behavior, related to their personal traits, interests, or recognition of stylized artwork.

In the first study, subjects were asked to freely explore three examples of classical artwork. Comparison of transition entropy H_t suggests that individuals who are more curious demonstrate significantly greater predictable patterns of visual scanning, resulting in lower H_t compared to those deemed less

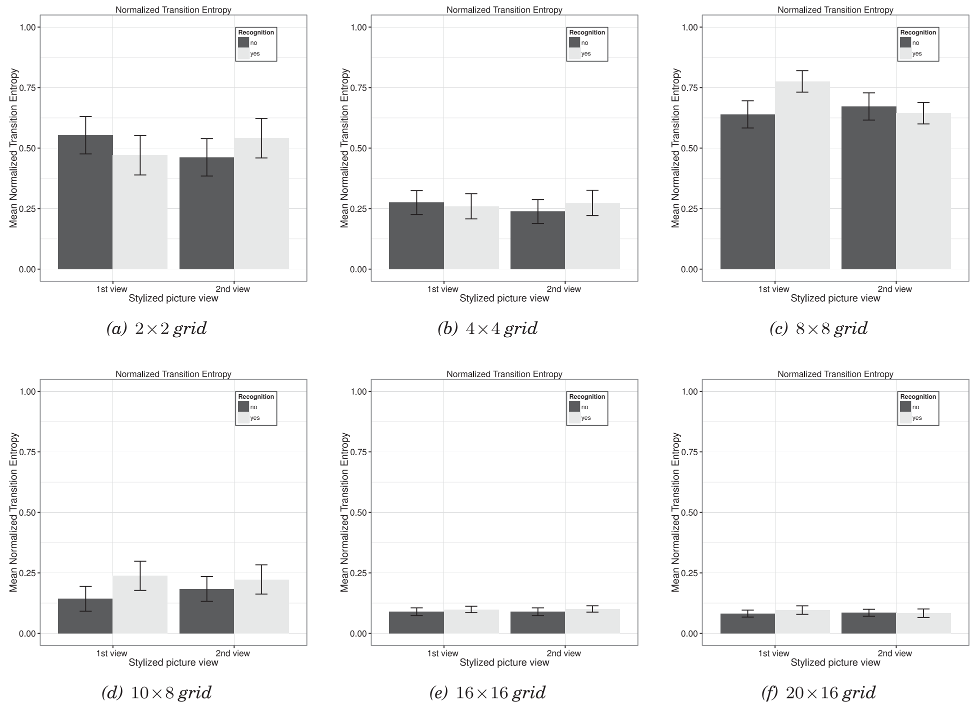


Fig. 9. Normalized transition entropies versus stimulus viewing task or versus recognition at different grid granularities. Error bars represent ±95% confidence intervals.

curious. Moreover, stationary entropy was related to individuals' subjective aesthetic evaluation of the artwork. Participants who better appreciated the artwork were characterized by lower stationary entropy H_s. We speculate that those who better appreciate the artwork devote their attention to preferred picture elements without exploring the remaining image with equal intensity.

In the second study, we demonstrated that higher transition H_t and stationary H_s entropies signify recognition of stylized classical art, namely *La Jaconde*. Those who recognized the painting in its stylized representation yielded higher stationary and transition entropies H_s and H_t than those who did not. In other words, they exhibited less predictable gaze switching behavior over the entire image. What is more, after presentation of the original *La Jaconde*, in the second viewing of stylized paintings, the differences between both stationary and transition entropies disappeared. In other words, after its presentation in its original form, the transition and stationary entropies of those who did not recognize the picture at first rivaled the entropies of those who recognized the picture in the first viewing.

Both presented studies suggest the transition and stationary entropies are potentially promising eye movement indicators of curiosity, interest, and picture familiarity. However, these relations cannot yet be treated as fully proven, and they need further evidence from hypothesis-driven experiments.

Future work should also address two issues related to the number and size of AOIs. More AOIs can lead to sparse matrices. Matrix sparsity should be further considered in terms of its influence on the interpretation of entropy. Future work is also needed to develop and improve the method of optimizing AOI grid size, probably based on the nature of the stimulus. The size of AOIs may also influence the interpretation of stationary entropy. Furthermore, AOI dynamics should also be considered. In principle, using fixated AOI scanpath indices (i.e., the *content-driven* analysis approach), the method could also be extended to dynamic stimuli provided that the set of AOIs $\mathcal{S} = \{1, \ldots, s\}$ (the state space) is fixed. The state space can operationally be fixed either over the entire exposure duration of the stimulus (e.g., scene vs. subtitle transitions when viewing subtitled film) or over a number of small temporal segments of the stimulus (e.g., constituting a frame-based approach as suggested by Grindinger et al. [2010a]). However, if the underlying stimulus changes dynamically, then the likelihood of fixating any given AOI is likely to change, which would affect the expected viewing probabilities over time. Thus, a *stimulus-driven* analysis would need to consider the dynamically changing fixation probabilities, as they are likely to change in response to the stimulus. For example, one could segment temporally more or less based on the constancy of the underlying stimulus, but this would be limited by the amount of fixations available per segment.

To conclude, modeling gaze transitions between AOIs as Markov chains and calculating normalized transition and stationary entropies produces straightforward measures of the predictability of saccadic transitions over stimuli. These measures can be tested for studying individual differences due to personal characteristics or stimuli properties via popular statistical tests.

ACKNOWLEDGMENTS

We would like to thank Dr. Cezary Biele and Ms. Agata Kopacz for sharing their data collected during the project on perception of artwork, which became an important part of this contribution.

REFERENCES

C. Acartürk and C. Habel. 2012. Eye tracking in multimodal comprehension of graphs. In *Proceedings of the Workshop on Technology-Enhanced Diagrams Research (TEDR'12)*. 11–25.

Robert Russell Althoff. 1998. *Eye Movement-Based Memory Assessment: The Use of Eye Movement Monitoring as an Indirect Measure of Memory*. Ph.D. Dissertation. University of Illinois at Urbana-Champaign.

Roman Bednarik, Niko Myller, Erkki Sutinen, and Markku Tukiainen. 2005. Applying eye-movement tracking to program visualization. In *Proceedings of the 2005 IEEE Symposium on Visual Languages and Human-Centric Computing (VLHCC'05)*. IEEE, Los Alamitos, CA, 302–304. DOI:http://dx.doi.org/10.1109/VLHCC.2005.20

Julian Besag and Debashis Mondal. 2013. Exact goodness-of-fit tests for Markov chains. *Biometrics* 69, 2, 488–496.

Cezary Biele, Agata Kopacz, and Krzysztof Krejtz. 2013. Shall we care about the user's feelings? Influence of affect and engagement on visual attention. In *Proceedings of the International Conference on Multimedia, Interaction, Design, and Innovation (MIDI'13)*. ACM, New York, NY, Article No. 7. DOI:http://dx.doi.org/10.1145/2500342.2500349

Gabriela Ciuperca and Valerie Girardin. 2005. On the estimation of the entropy rate of finite Markov chains. In *Proceedings of the 11th International Symposium on Applied Stochastic Models and Data Analysis*.

Thomas M. Cover and Joy A. Thomas. 2006. *Elements of Information Theory* (2nd ed.). John Wiley & Sons.

Jhani A. de Bruin, Katherine M. Malan, and Jan H. P. Eloff. 2013. Saccade deviation indicators for automated eye tracking analysis. In *Proceedings of the 2013 Conference on Eye Tracking South Africa (ETSA'13)*. ACM, New York, NY, 47–54. DOI:http://dx.doi.org/10.1145/2509315.2509324

Andrew T. Duchowski, Sabarish V. Babu, Jeff Bertrand, and Krzysztof Krejtz. 2014. Gaze analytics pipeline for unity 3D integration: Signal filtering and analysis. In *Proceedings of the 2nd International Workshop on Eye Tracking for Spatial Research (ET4S'14)*.

Andrew T. Duchowski, Jason Driver, Sheriff Jolaoso, William Tan, Beverly N. Ramey, and Ami Robbins. 2010. Scanpath comparison revisited. In *Proceedings of the 2010 Symposium on Eye-Tracking Research and Applications (ETRA'10)*. ACM, New York, NY, 219–226. DOI:http://dx.doi.org/10.1145/1743666.1743719

Laura Ekroot and Thomas M. Cover. 1993. The entropy of Markov trajectories. *IEEE Transactions on Information Theory* 39, 4, 1418–1421.

Stephen R. Ellis and Lawrence Stark. 1986. Statistical dependency in visual scanning. *Human Factors* 28, 4, 421–438.

Miquel Feixas, Esteve del Acebo, Philippe Bekaert, and Mateu Sbert. 1999. An information theory framework for the analysis of scene complexity. In *EuroGraphics*, P. Brunet and R. Scopigno (Eds.), Vol. 18. EuroGraphics, 95–106.

Peter Fischer and Elizabeth Peinsipp-Byma. 2007. Eye tracking for objective usability evaluation. In *Proceedings of the European Conference on Eye Movements (ECEM'07)*.

Joseph H. Goldberg and Xerxes P. Kotval. 1999. Computer interface evaluation using eye movements: Methods and constructs. *International Journal of Industrial Ergonomics* 24, 631–645.

Thomas Grindinger, Andrew T. Duchowski, and Michael Sawyer. 2010a. Group-wise similarity and classification of aggregate scanpaths. In *Proceedings of the 2010 Symposium on Eye Tracking Research and Applications (ETRA'10)*. ACM, New York, NY, 101–104.

Thomas J. Grindinger, Vidya N. Murali, Stephen Tetreault, Andrew T. Duchowski, Stan T. Birchfield, and Pilar Orero. 2010b. Algorithm for discriminating aggregate gaze points: Comparison with salient regions-of-interest. In *Proceedings of the International Workshop on Gaze Sensing and Interactions (IWCGI'10)*. 390–399.

Jean Hausser and Korbinian Strimmer. 2009. Entropy inference and the James-Stein estimator, with application to nonlinear gene association networks. *Journal of Machine Learning Research* 1469–1484. http://jmlr.org/papers/v10/hausser09a.html.

Elizabeth Hume and Frédéric Mailhot. 2013. The role of entropy and surprisal in phonologization and language change. In *Origins of Sound Change: Approaches to Phonologization*, Alan C. L. Yu (Ed.). Oxford University Press, Oxford, UK, 29–47.

Alex D. Hwang, Hsueh-Cheng Wang, and Marc Pomplun. 2011. Semantic guidance of eye movements in real-world scenes. *Vision Research* 51, 1192–1205.

Halszka Jarodzka, Kenneth Holmqvist, and Marcus Nyström. 2010. A vector-based, multidimensional scanpath similarity measure. In *Proceedings of the 2010 Symposium on Eye-Tracking Research and Applications (ETRA'10)*. ACM, New York, NY, 211–218. DOI:http://dx.doi.org/10.1145/1743666.1743718

Todd B. Kashdan, Matt W. Gallagher, Paul J. Silvia, Beate P. Winterstein, William E. Breen, Daniel Terhar, and Michael F. Steger. 2009. The curiosity and exploration inventory-II: Development, factor structure, and psychometrics. *Journal of Research in Personality* 43, 987–998.

Krzysztof Krejtz, Tomasz Szmidt, Andrew T. Duchowski, and Izabela Krejtz. 2014. Entropy-based statistical analysis of eye movement transitions. In *Proceedings of the 2014 Symposium on Eye-Tracking Research and Applications (ETRA'14)*. ACM, New York, NY. DOI:http://dx.doi.org/10.1145/2578153.2578168

Anje Kruizinga, Ben Mulder, and Dick de Waard. 2006. Eye scan patterns in a simulated ambulance dispatchers task. In *Developments in Human Factors in Transportation, Design, and Evaluation*. Shaker Publishing, Orbassano, Italy.

Solomon Kullback and R. A. Leibler. 1951. On information and sufficiency. *Annals of Mathematical Statistics* 2222, 79–86.

John Liechty, Rik Pieters, and Michel Wedel. 2003. Global and local covert visual attention: Evidence from a Bayesian hidden Markov model. *Psychometrika* 68, 4, 519–541.

Vincente Ponsoda, D. Scott, and J. M. Findlay. 1995. A probability vector and transition matrix analysis of eye movements during visual search. *Acta Psychologica* 88, 167–185.

R Development Core Team. 2011. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/ ISBN 3-900051-07-0.

Frederick Shic, Katarzyna Chawarska, Jessica, and Brian Scassellati. 2008. Autism, eye-tracking, entropy. In *Proceedings of the 7th IEEE International Conference on Development and Learning (ICDL'08)*. 73–78.

Lisa Vandeberg, Samantha Bouwmeester, Bruno R. Bocanegra, and Rolf A. Zwaan. 2013. Detecting cognitive interactions through eye movement transitions. *Journal of Memory and Language* 69, 445–460.

Boris M. Velichkovsky, Markus Joos, Jens R. Helmert, and Sebastian Pannasch. 2005. Two visual systems and their eye movements: Evidence from static and dynamic scene perception. In *Proceedings of the 27th Conference of the Cognitive Science Society (CogSci'05)*. 2283–2288.

Robin S. Weiss, Roger Remington, and Stephen R. Ellis. 1989. Sampling distributions of the entropy in visual scanning. *Behavior Research Methods, Instruments, and Computers* 21, 3, 348–352.

Received September 2014; revised September 2015; accepted September 2015