Veronica Alampi Sottini[1], Elena Barbierato[1], Iacopo Bernetti[1], Irene Capecchi[1], Maria Cipollaro[1], Sandro Sacchelli[1], Claudio Saragosa[2]

Urban landscape assessment: a perceptual approach combining virtual reality and crowdsourced photo geodata

[1] *Department of Agricultural, Food and Forestry System - GESAAF, Università degli Studi di Firenze*
[2] *Department of Architecture – DIDA, Università degli Studi di Firenze*

E-mail: *veronica.alampi@unifi.it, elena.barbierato@unifi.it, iacopo.bernetti@unifi.it, irene.capecchi@unifi.it, maria.cipollaro@unifi.it, sandro.sacchelli@unifi.it, claudio.saragosa@unifi.it*

The paper focuses on the evaluation of the visual quality of the cityscape of Livorno, in Italy. Using an algorithm based on Flickr's Application Programming Interface, the coordinates of 7,453 shooting points of a shared photo were downloaded. A direct survey was conducted to assess the perception of urban landscape. The geographical spaces to be evaluated were measured using 78 spherical images played in virtual reality through a virtual reality headset.

The data were processed with the Principal Components Analysis. A Visual Quality Index (VQI) was developed to spatialise the results. Saliency maps were calculated to survey typical patterns in areas with high VQI. The results demonstrate how subjective visual perception helps identifying the criteria for the design of high-quality public spaces.

1. Introduction

The contemporary and modern town is the result of the rationalist city characterized by a functionalist concept that overshadowed the aesthetic dimension. It has defined spaces which are increasingly less welcoming and in which the individual can no longer recognize themself or establish an empathetic relationship. Settis (2017) suggests the term "community's dysmorphophobia", to connote the phenomena of "loss of collective image", using an expression coined on the pathologies of the individual's mind. Lynch's pioneering researches (1960) investigate the possible links between psychology (specifically on the perception of shape) and the configuration of the urban environment, thereby trying to understand which elements give the space in which we are immersed those qualities that make it vivid, recognizable and meaningful. To study these themes, Lynch (1960) proposes a perceptive survey with open questions that focus on the sense of orientation of the respondent and on the relationships between the individual and the surrounding environment. Using questionnaires, Lynch (1960) defines the cognitive map that each community creates for its own city and sets a public rep-

resentation or a collective image of it, by analysing the most recurrent elements. He can identify which elements and characteristics are attractive to people and determine the perceptive quality of the urban environment. What Lynch (1960) calls "imageability". The higher the "imageability", the more likely the urban environment is to evoke an image that impresses the observer's mind. Lynch also introduces the degree of "legibility", i.e. "the ease through which the parts of the image can be recognised and organised in a coherent system".

Lynch (1960), referring to the fundamental studies of Gestalt theory, traces a path of research that continues until today, thanks to the studies related to perception from the world of neuroscience. The process, reconstructed by Mallgrave (2015) and Saragosa (2016), tries to explore how the emotion produced by the image of the environment around us favours processes of well-being (or discomfort, dysmorphophobia). The path of research leads to the current debate on mirror neurons and the resumption of the theory of empathy (better known today as embodied simulation), a path that leads us to think differently about the space that welcomes us, which is therefore not only phenomenologically different (anisotropic, as we have seen), but also perceptively dissimilar. In other words, not everything that is configured around us can produce the same emotional quality, the same empathic vibration.

Our research tries to discover new tools to evaluate which configurations of the urban space contribute to stimulating the emotional aspects of the inhabitants. If the research could provide clarity in this direction, it would become important for the definition of useful tools for urban regeneration design, one of the fundamental objectives of the 21st century urban planning. Studies on urban configuration will probably help to overcome the cacotopias (Geddes, 1971; Mumford, 2017) of the anthropic environment that has been produced over time. As for how much the environment in which we are immersed awakens our emotions and affects our lives, Fitzgerald's vivid description in his novel "Tender is the Night" (2001) illustrates it beautifully when he describes the Divers in the sun of southern France:

> But the diffused magic of the hot sweet South had withdrawn into them – the soft-pawed night and the ghostly wash of the Mediterranean far below – the magic left these things and melted into the two Divers and became part of them.

The assessment of the visual quality of space has been addressed mainly regarding landscape assessment (Tempesta, 2009). In this context, two methods have been developed in the evaluation of the visual quality of the space around us. These two approaches have produced specific methodologies: the objectivist (or physical) and the subjectivist (or psychological) paradigm (Lothian, 1999). In the case of the objectivist approach, the quality of the living space would derive from its formal components and the combination of these components allows the assessment of its quality. The subjectivist approach focuses on instinctive, emotional and experiential aspects, and the quality of space. The assessment process must first analyse and understand the reactions of a man placed in different environments and subsequently it creates the rules that combine the formal elements.

From an operative point of view, the subjectivist paradigm is based on a questionnaire survey on the perceptive evaluations of a sample of individuals who occasionally or regularly interact with the space under study (de Val, Atauri and de Lucio, 2006; Zube, Sell and Taylor, 1982).

The most recent techniques in this field make use of surveys and virtual reality technologies (for detecting reactions to stimulus) and consider the semantic differential method (Brown, de Bie and Weber, 2015; Kuliga *et al.*, 2015; Lim, Honjo and Umeki, 2006).

Recently, with the development of Internet social networking platforms (the so-called social media) several approaches have been developed that work on the use of big data (Jin *et al.*, 2010) made available by social media platforms. An implementation in the city of San Francisco highlights (Dunkel, 2015) the great potential of photographic data shared by the Flickr platform for the analysis of environmental perception in landscape and urban planning. Flickr is an important source of additional information for a perspective on the perception of the urban environment by the general public. Zhou, Xu and Kimmons (2015) automate the detection of places of interest in multiple cities based on spatial and temporal features of Flickr images from 2007 on. Hauthal and Burghardt (2016) used Flickr data for the city of Dresden (Germany) to extract the emotions related to the perception of the urban space through a sentiment analysis of the comments associated with the city. The study shows the complexity of the Shared Geographical Information, which allows the extraction of information about the user regarding space and time. The study concludes by saying that categorizing users into user groups to analyse their behaviour can replace (or at least complete) empirical investigations.

The most critical aspects of the methods are the following:
- sampling the space to identify the representative visual stimuli (urban or rural landscape views);
- identifying methods of providing the visual stimulus of the space under evaluation to the respondent;
- the setting of the questionnaire that should allow the emotional reaction of the respondent to be detected.

For the first aspect, the most recent techniques use the voluntary geographical information from social media big data sets and the technologies of survey and virtual reality. Semantic differential techniques are used for the detection of the reactions to the stimulus (Chen and Takamatsu, 2016; Green, 1999; Lim, Honjo and Umeki, 2006).

The usefulness of identifying urban spaces with higher visual quality lies in the possibility of searching in these areas specific rules able to replicate positive emotions to be used then for the design and regeneration of urban spaces. Following Pallasmaa *et al.* (2015, p. 5), "Contemporary architecture has often been accused of emotional coldness, restrictive aesthetics and a distance from life. This criticism suggests that we architects have adopted formalist attitudes, instead of tuning our buildings with realities of life and the human mind." In the work "A

pattern language" Alexander, Ishikawa and Silverstein (1977) hypothesise the existence of specific rules of composition of the space (the patterns) that stratify over time and that follow specific morphotypological laws, generating urban quality.

On the basis of the above, the aims of the paper are:

- to assess the quality of urban space by applying a subjectivist approach through a questionnaire based on eliciting the perception of individuals. This is carried out implementing a semantic differential approach combined with virtual reality techniques;
- to spatialise the relationship between the quality of the urban space assessed through sampling and the sharing of voluntary geographical information;
- to survey typical patterns in areas with high visual quality to define urban design and regeneration rules.

The results will be further used for the determination of objective models for the improvement of the existing and the design of new open public spaces in the city. Among the various dimensions of urban design (Carmona *et al.*, 2012), the physical one and the aesthetic-visual criterion play a primary role in the process of designing the space with recognizable socio-cultural values and identities. Several theoretical works (Arnheim, 1969; Asihara, 1983; Gibson, 1979) demonstrated the importance of developing multidimensional perceptive analyses for the development of quality urban spaces (Perovic and Folic, 2012).

2. The method

2.1 The study area

Livorno is a municipality of Central Italy located on the Tuscan coast and overlooking the Ligurian Sea. The municipality covers an area of 104.5 square kilometres and counts 157,052 inhabitants. Livorno is the third largest municipality in Tuscany by population. The perimeter of the study area only includes the urbanized territory of the city of Livorno, since the topic of the paper focuses on urban quality. Thus, industrial areas to the north of the city have been excluded from the municipal territory, as for some neighbouring areas that are too far from the city centre and those that are on the border with rural areas.

2.2 The assessment of the quality of urban space through shared geographical information

In the absence of formal reference structures, geographical information generated by web users is in some respects chaotic: flows of information are created and move in different directions, crossing and exchanging contents, while losing any separation between production and consumption and justifying the widespread notions of "produser" (Bruns, 2006) and "produsage". Over time, despite the disorganization and heterogeneity that characterize this information, its pro-

Figure 1. Flow-chart of the proposed approach.
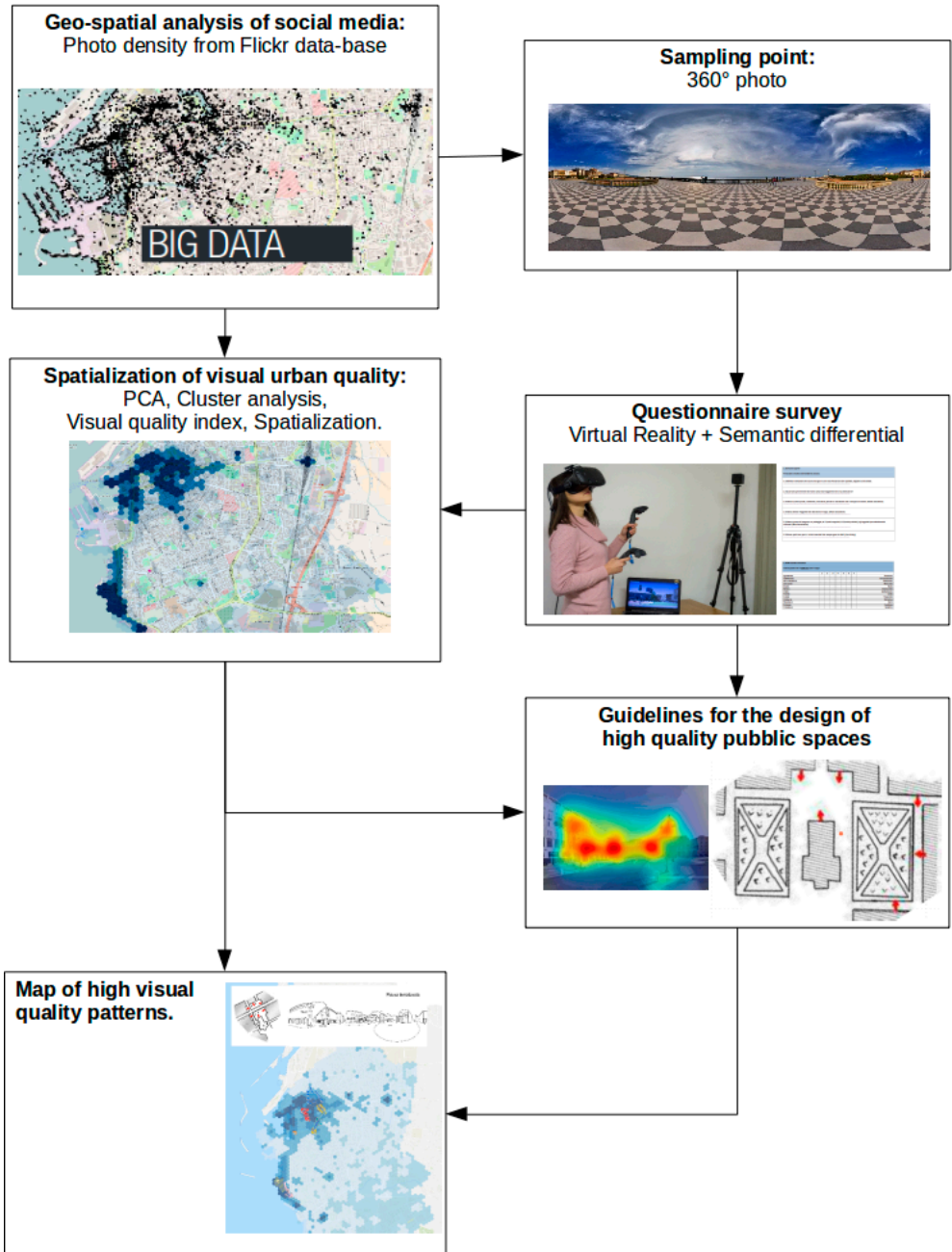
duction has become such an extensive phenomenon attracting the interest of the scientific community and of people historically engaged in the creation and distribution of geographical information (Borruso, 2013).

We are currently experiencing a rapid increase in available data sources regarding voluntary geographical information. Social media applications such as Twitter, Flickr or Facebook provide a collective of geographical information that can be interrogated via public Programming Interfaces (APIs). In addition, we are witnessing a growing willingness of people to actively contribute by sharing their experiences of living the urban, rural and natural space in a context of use that falls under the broad term "People as sensors".

In recent years many researchers used big data from social media platform to analyse the users' preference in various fields. There are a lot of social media platforms, the most used of which is Facebook, but the data are not easily accessible. Other platforms give an easier access to public data, like Flickr, Twitter, Instagram. Flickr (founded in 2004) is used by professional or amateur photographers to share photos, Twitter (founded in 2006) is used to share short text messages, while Instagram (foundedin 2010) is used to share photos, usually related to everyday life, taken by smartphones with high quality cameras. In research studies, Twitter is preferably used when dealing with business and communication analysis, Flickr in the environmental sciences (Spyrou and Mylonas, 2016), while Instagram is an unexplored source of data, even though it is becoming popular (Tenkanen *et al.*, 2017). Thus, we decided to use Flickr, not only because it is the most suitable for environmental sciences studies, but also it is exclusively used by professional or amateur photographers because it was the first social media platforms to be created and, today, it has a photographic database size larger than Instagram.

According to Nov, Naaman and Ye (2010), the photographic data uploaded on the Flickr platform imply an individual process that can be divided into two main phases:
a) the technical-creative phase of taking the photo;
b) the social phase of sharing this photo by associating commentary information to it.

Lynch (1960: p. 4) suggests that

[. . .] the generalized mental picture of the exterior physical world that is held by an individual [. . .] is the product both of immediate sensation and of the memory of past experience, and it is used to interpret information and to guide action.

Speaking generally (Collier, 1967; Dakin, 2003; Scott and Canter, 1997; Sontag, 1977), the action of taking a picture is not only linked to the characteristics of the surrounding environment, but involves all of the aspects of the interpretative cognition that the individual applies to that space (personal preferences, memories, opinions, etc.). So, both the act of taking a picture in a specific place and the consequent action of choosing which photos to share on the social network platform reflect the quality of the perception that the individual has of that place.

There are many types software able to map and evaluate the ecosystem services using the Flickr platform, such as the software model InVEST. The InVEST recreation model calculates the total number of photo-user-days for each grid cell or polygon. The tool, which uses a simple linear regression, compares photo-user-day with the chosen landscape index to analyse how each attribute influences the visitation rate. The results are maps that identify current and future patterns of recreational use.

We decided to use an algorithm based on Flickr's Application Programming Interface instead of other types of software, to have a major flexibility both in the method and in the analysis of the results.

We downloaded the coordinates of 7,453 shooting points of shared photo uploaded from 2005 to 2017. The number of photos uploaded to Flickr showed a peak in the years 2011-2012, it decreased in subsequent years, but it remained for research purposes. We assumed that the preferences of the individual did not change in the period 2011- 2017.

The point data has been transformed into a density surface by an analysis of Kernel Density Estimation (KDE) (Chen and Shaw, 2016).

Kernel Density Estimation (KDE) is one of the most popular methods in geostatistical analysis. It uses a continuous function in the space to define density. This approach can be followed drawing a circle of radius r around each point and dividing point number inside the circle by its area. KDE is a moving three-dimensional function and allows to weight the points within the circle, assigning a higher value to the points closest to the centre of the circle, and a lower value to the points closest to the circumference (Danese, Lazzari and Murgante, 2008).

The only arbitrary variable in the KDE is represented by the bandwidth (Bailey and Gatrell, 1995). Different values of bandwidth produce different patterns: if the value of bandwidth is high, the distribution tends to be smoother, while if the value of bandwidth is low there will be more peaks in the distribution.

In the case under study, a triangular Kernel with a radius of 50 metres has been used. The kernel was then overlapped with a 100-metre side hexagonal grid. In the case under study, a triangular Kernel with a radius of 50 metres has been used. The kernel was then overlapped with a 100-metre side hexagonal grid. The size and the shape were chosen because of their topological and geometric properties (Feick and Robertson, 2015; Patil *et al.*, 2000; White, Kimerling and Overton, 1992).

Looking at the city of Livorno, it is possible to notice an uneven distribution of the points. Specific areas are much more photographed than others. In particular, the most photographed places are the areas of Venezia Nuova and Terrazza Mascagni.

2.3 The assessment of the visual quality of urban space by sampling

The definition of the experimental design required the creation of two sampling schemes: a spatial/geographical type, to choose the locations of the city of

Livorno in which to detect the stimuli to be submitted to the interviewees; the definition of the questionnaire and the choice of the interviewees.

For the spatial sampling, the areas to be surveyed have been chosen so as to consider the interactions between individuals and urban space using shared geographical information categorised and described in the previous paragraph. The photo density map was overlapped with a hexagonal grid with the side of 100 meters. The sampling points were identified by means of a stratified random sampling by extracting the cells from that grid. In particular, the cells of the grid were divided into three layers: high density photo cells, beyond the 90th percentile; medium density photo cells, between the 90th and the 50th percentile; low density photo cells, lower than the 50th percentile. The sample layers were chosen in such a way as to over-sample the areas with a high concentration of photos, as they are of greater interest for the purpose of the study. Lastly, 26 random points were extracted in each layer, for a total of 78 points. On these points the stimuli to undergo the assessment were detected. The geographical spaces to be evaluated were measured using 360° spherical images downloaded from the Google Street View database. Those images can be played in virtual reality through a virtual reality headset. As highlighted by Piga's research (Piga 2011, 2013), the different types of photographic reproduction media of urban and landscape space – panoramic photos, videos or 360° photos – allow many different types of interaction:
- panoramic images provide a unique point of view (specific visualization)
- videos reproduce the visualisation of movement in space (moving visualization).

In the first case the observers can decide where to look, as if they were inside an infinite fixed moment, while in the second they pass through the space. However, in general the observers cannot choose where to point their gaze. In comparison to the first two cases, the use of 360° interactive photographs and videos through the virtual reality headset allows an interactive and immersive visualization of the space that surrounds the interviewee. In all the above cases the point of view is that of the camera and the observer is not free to roam; only virtual reality gives this possibility, approximating to the real experience in space (Piga, 2011, 2013). For this reason, the quantitative relationships between the experience of the architectural spaces and the physical properties were studied using virtual reality simulation (Franz *et al.*, 2005; Lovett *et al.*, 2015; Maffei *et al.*, 2015; Orland, Budthimedhee and Usitalo, 2001; Orenstein, Zimroni and Eizengerg, 2015).

The elicitation of the perceptions was carried out by means of a questionnaire administered through the virtual reality headset to a sample of 234 individuals. Participants were selected using random sampling techniques. The interviewees had to be:
(i)   local people, i.e. residents in the city of Livorno
(ii) non-local people, i.e. undergraduate students attending in urban planning disciplines at the University of Florence, or alternatively
(iii) residents in another province or territory in Tuscany.

For each sample layer, 78 individuals were interviewed, each of them were given two questionnaires by randomly choosing the spatial sample layer (low, medium or high density photos).

The questionnaire was structured into two sections. The first one aimed at collecting the personal data of the interviewees, through open questions on age, city of residence, field and level of education.

The second part of the questionnaire was structured as to detect the perception of the quality of urban space through the virtual reality headset. To detect the individual subjective perception of the interviewees, the survey was based on the semantic differential technique (Osgood, *et al.*, 1957) measured on a 7-point Likert scale. To evaluate both the perceptive and structural emotional aspect of the examined place, two series of bipolar semantic couples were used: one made of bipolar adjectives inherent to the emotional sphere, the other of opposite adjectives inherent to the structural and physical aspects of the stimulus. The pairs used are shown in Table 1. The terms used were arranged alternately: a bipolar pair is followed by a structural pair. The choice of the semantic couples was based on the works of Bonaiuto *et al.* (2003), Green (1999), Hung and Nieh (2009), Hsu *et al.* (1999), Imamoglu (2000), Kang and Zhang (2010). The adjectives linked to the emotional sphere are in a smaller number than the structural ones, because only those pairs that were found to have a more diversified meaning were chosen among the various adjectives pairs available in literature.

During the administration of the questionnaire, a video was recorded to detect the direction of the interviewee's sight in the virtual reality headset while observ-

Table 1. Bipolar pairs.

Emotional Adjectives	Formal Adjectives
Unpleasant- Pleasant	Narrow– Wide
Indifferent – Exciting	Homogeneous–Different
Non Stimulating – Stimulating	Simple –Complex
Stressful – Relaxing	Asymmetric–Symmetric
Ugly– Beautiful	Plain–Decorated
Boring – Interesting	Closed– Open
Unnerving – Comfortable	Monotone –Varied
	Irregular– Regular
	Disordered–Ordered
	Ordinary– Special
	Uncared for - Cared for
	Dynamic–Static
	Monochromatic– Coloured

Source: our elaboration.

ing the spherical photo. An example of this video is available in the supplementary materials relative to sampling point n. 31.

2.4 Statistical analysis.

The most suitable methodology for the statistical elaboration of the data collected through the questionnaires was identified in the Principal Component Analysis (PCA).

The PCA is a data synthesis technique, developed by several researchers, including Pearson (1901), Cauchy (1829), Jordan (1874) and Hotelling (1933), the latter having defined the term and the method (Stewart, 1993).

The PCA's primary goal is to simplify data by reducing the number of starting variables, correlated among them, in a number of minor independent variables, that are called latent variables or principal components. This technique is one of the most used, because it reduces the number of variables, revealing new and more evident relationships that had not initially been considered.

Moreover, while choosing the appropriate number of variables, it is possible to monitor the "trade-off" between information loss and simplification of the problem. (Abdi, Lynne, Williams, 2010).

The second step in the statistical analysis considered the use of the principal components to create a synthetic Visual Quality Index (VQI). Di Franco and Marradi (2003) proposed the methodology. The constructed index allows attributing a value (score) to each subject (in this case, sites subjected to photographic sampling). The creation of an index that detects the phenomenon described by one or more principal components, obtained by analysing a set of variables with the PCA technique, was carried out according to the following procedure. The method is based on the inspection of the two-dimensional projections of space and the consequent isolation and separate refinement of each component extracted, thus creating a PCA in two distinct stages. The purpose of the first PCA is to identify the dimensions underlying the set of variables analysed. Once decided how many components to extract, for each of them there is a cluster of variables (subset of the source set) that are strongly correlated with each other and with their relative size. At the end of this stage, the original set of variables is divided into a number of subsets, equal to the number of components to be extracted. For each identified subset, subsequent PCAs are carried out, progressively eliminating the variables that are less correlated to the relative dimension, until the component is refined: i.e., only the variables strongly correlated with the dimension will remain. Only the selected variables are used for the creation of the index. The calculation of the component coefficients is carried out by performing the PCA only on the subset of the final variables obtained from the refinement process.

The final step is to use the component coefficients to calculate a final indicator for each observation (sample point) according to the formula:

$$VQIC_{,i} = \sum z_{j,i} \cdot C_{c,j}.$$ [1]

with:
VQIC,i = Visual Quality Index for the Component C=I, II and the sample point i = 1, ... 78
j = (sub)set of the semantic differential variables (the whole subset for method 2 and the residual variables after refinement for method 1)
$c_{C,j}$ = component coefficient of variable j on component C
$z_{j,i}$ = standardised scores obtained on the semantic differential variable j, for each sample point i.
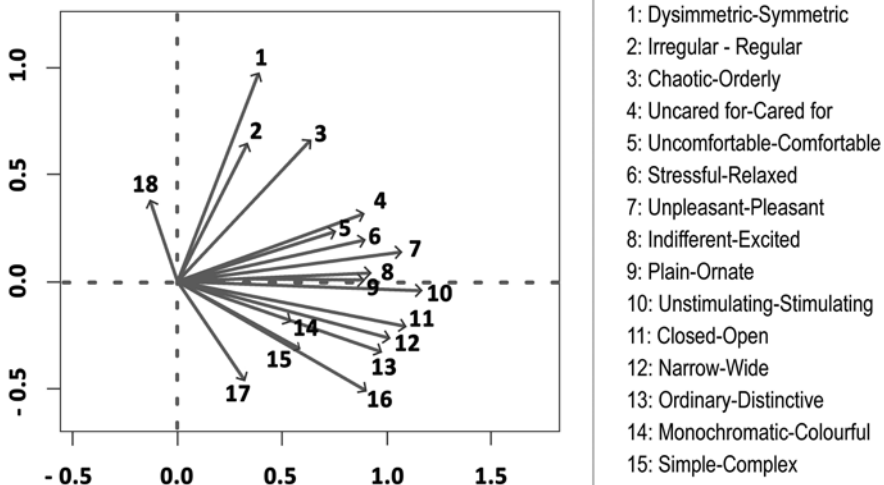
3. Result

In this study, the analysis of the variance explained by the first dimensions shows the presence of a strong relationship among the variables and suggests the number of dimensions that should be considered. The first two dimensions of PCA represent 64.99% of the total inertia data; this means that 64.99% of the complex cloud variables are explained by the space defined by the two extracted dimensions. This percentage is relatively high and therefore the space defined by the two components, where the cloud of the variables is projected, well represents the variability of the data. This value is significantly higher than the reference value of 22.2%, obtained through Parallel Analysis.

The map of variables (Graph 1) has made it possible to define the starting variables that characterize and define the first two dimensions examined.

Those variables that are positively and exclusively correlated to the first dimension are: Unpleasant-Pleasant, Narrow-Wide, Indifferent-Exciting, Non Stimulat-

Graph 1. Map of variables.

1: Dysimmetric-Symmetric
2: Irregular - Regular
3: Chaotic-Orderly
4: Uncared for-Cared for
5: Uncomfortable-Comfortable
6: Stressful-Relaxed
7: Unpleasant-Pleasant
8: Indifferent-Excited
9: Plain-Ornate
10: Unstimulating-Stimulating
11: Closed-Open
12: Narrow-Wide
13: Ordinary-Distinctive
14: Monochromatic-Colourful
15: Simple-Complex

ing-Stimulating, Simple-Complex, Stressful-Relaxing, Plain-Decorated, Unnerving-Comfortable, Closed-Open, Ordinary-Special, Uncared for-Cared, Monochromatic-Coloured. It is possible to notice that there is a correlation between the emotional variables (Unpleasant-Pleasant, Indifferent-Exciting, Non stimulating-Stimulating, Stressful-Relaxing, Ugly-Beautiful, Boring-Interesting, Unnerving-Comfortable) and the structural ones (Narrow-Wide, Homogeneous-Different, Simple-Complex, Asymmetric-Symmetric, Plain-Decorated, Closed-Open, Monotone-Varied, Irregular-Regular, Disordered-Ordered, Ordinary-Special, Uncared for-Cared for, Dynamic-Static, Monochromatic-Coloured). Thus, if a place has been rated pleasant, beautiful, relaxing, it shows the formal characteristics of cared for, wide, varied, special, open, decorated. In summary, the first dimension appears correlated as a whole to the emotional perception of the quality of urban space.

On the contrary, the second dimension is positively saturated by Irregular-Regular, Asymmetric-Symmetric and negatively correlated by Homogeneous-Different and it can therefore be interpreted as correlated to the perception of some formal characteristics of the urban environment.

The pair of adjectives Monotone-Varied is positively correlated to the first dimension and negatively to the second, while the item Ordinary-Special is positively correlated to both. The pair of adjectives Dynamic-Static is not correlated to any of the two dimensions, thus highlighting the insignificance in the description of the phenomenon.

The results of the VQI, obtained for the two components (first component "emotional" and second component "formal") are shown in Table 2, Fig. 2 and Graph 2.

Fig. 2 indicates how only the first component "emotional" records higher total VQI in the area with a high photo density compared to that obtained in the area with medium and low density. This result is confirmed by the frequency distribution and the relative significance of difference among the averages of the values shown in Graph 2. The first component shows higher average values in the area with a high photo density compared to medium and low-density areas with a high significance. There are no significant differences in the values obtained for low and medium density areas. The Visual Quality Index for the second component "formal" does not discriminate among the different areas in either of the two methods.

In summary, the results obtained show that the photo density detected by the Flickr platform is correlated to a high visual quality deriving from the semantic differentials mostly related to the instinctive dimensions only for high values of photo density. It must be highlighted that the high Flickr photo density area represents the sampling hexagons with values beyond the 90th percentile of the entire sampled area.

A regression analysis was then performed to investigate the relationship between the visual quality revealed by the questionnaires and the geographical photo density downloaded from the Flickr platform. Since the relationship investigated is highly non-linear, it was not possible to adopt a parametric estimation method. The technique chosen was the regression tree.

Table 2. Results of the components' refinement.

Variable	Variable type	PCA I First Dimension "emotional"	PCA II Second Dimension "formal"	PCA I - coefficients before refinement process	PCA II- coefficients before refinement process
Unpleasant–Pleasant	Emotional Adjectives	0.9229	0.1323	0.9183	-
Non Stimulating-Stimulating	Emotional Adjectives	0.9149	-0.0183	0.9161	-
Indifferent–Exciting	EmotionalAdjectives	0.8649	0.0313	0.8666	-
Closed – Open	Formal Adjectives	0.8149	-0.1806	0.8356	-
Plain–Decorated	Formal Adjectives	0.8343	0.0349	0.8208	-
Stressful–Relaxing	EmotionalAdjectives	0.8113	0.1774	0.8130	-
Monotone –Varied	Formal Adjectives	0.7868	-0.4330	0.8027	-
Narrow–Wide	Formal Adjectives	0.7727	-0.2052	0.7995	-
Ordinary–Special	Formal Adjectives	0.7904	-0.2611	0.7928	-
Unnerving–Comfortable	Emotional Adjectives	0.7974	0.2581	0.7925	-
Uncared for- Cared for	Formal Adjectives	0.7703	0.2674	0.7472	-
Monochromatic– Coloured	Formal Adjectives	0.5562	-0.2116	-	-
Simple –Complex	Formal Adjectives	0.6278	-0.3596	-	-
Asymmetric–Symmetric	Formal Adjectives	0.3199	0.8078	-	0.7494
Irregular– Regular	Formal Adjectives	0.3824	0.7088	-	0.6467
Disordered–Ordered	Formal Adjectives	0.6130	0.6070	-	0.5195
Dynamic–Static	Formal Adjectives	-0.1378	0.4028	-	0.4374
Homogeneous–Different	Formal Adjectives	0.3467	-0.5126	-	-0.5785

Source: our elaboration.

Figure 2. Visual Quality Index of the sample points.
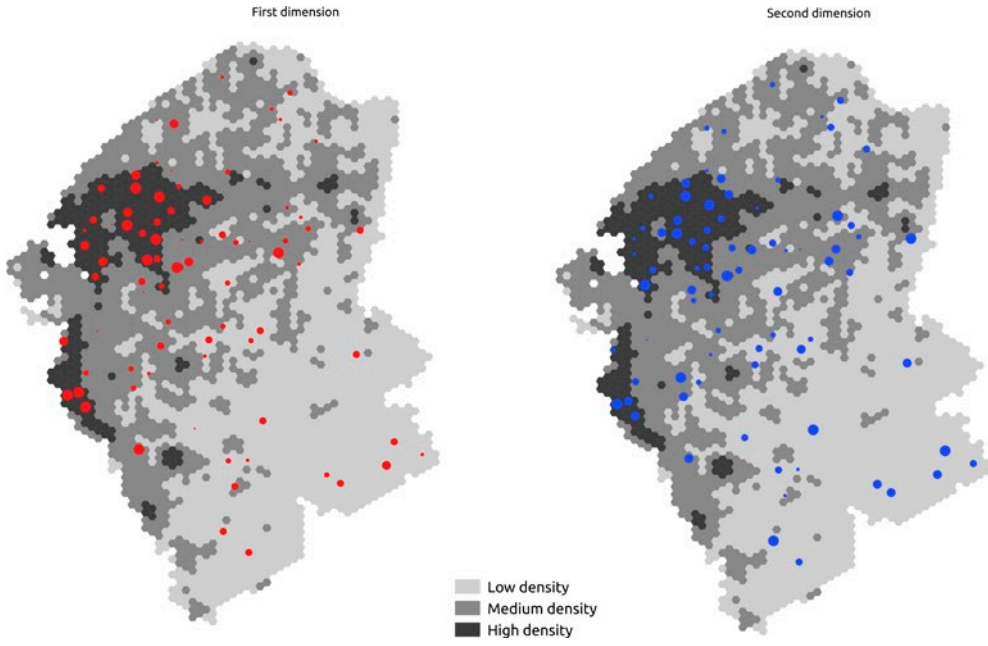
Graph 2. Frequency distribution of visual quality indices.
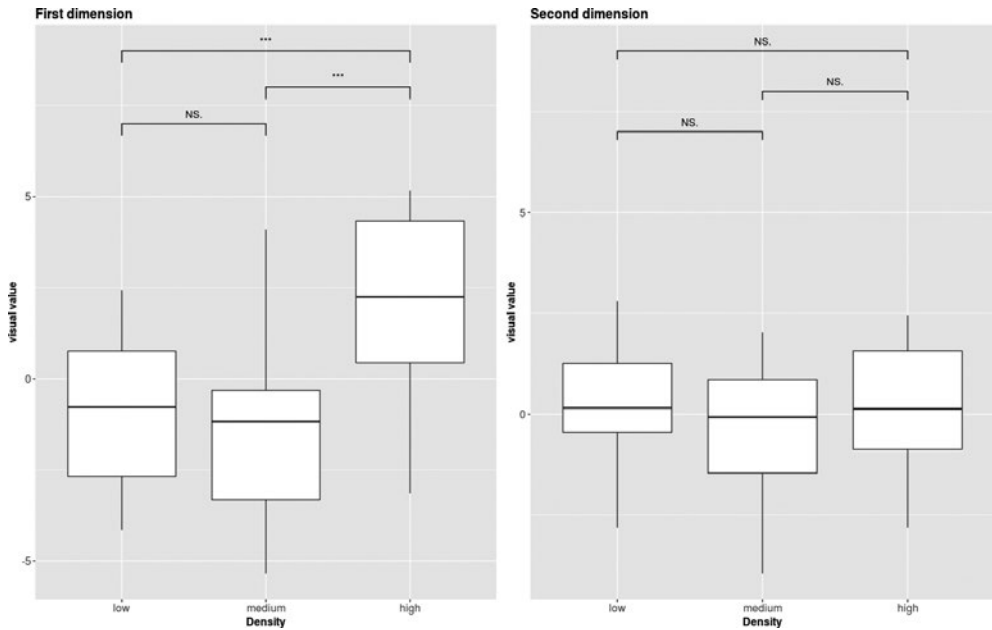
Graph 3 shows the diagram of the regression tree obtained through the estimation calculated on the Visual Quality Index. The photo density at which the terminal nodes of the classification originated was as follows, in an ascending order of photo density:

- Leaf 6 with a photo density of less than 0.084 (25th percentile) and a VQI of -0.2922;
- Leaf 4 with a photo density between 0.084 and 3.54 (25th and 82nd percentile) and a VQI of -0.18;
- Leaf 5 with a photo density between 3.54 and 12.49 (82nd and 92nd percentile) and a VQI of -0.20;
- Leaf 8 with a photo density between 12.49 and 23.87 (92nd and 97th percentile) and a VQI of 1.66;
- Leaf 9 with a photo density greater than 23.87 (beyond the 97th percentile) and a VQI of 3.41.

The normalized mean squared error was calculated for the evaluation of the predictive capacity of the model adopted (Torgo, 2016). The result was 0.73, a value to be considered satisfactory. The results confirm the difference among the averages of the sampled areas and broaden the information obtained by highlighting how the photo density is an efficient estimator of the perceived quality of places, especially when the photographs of the spaces are shared on social networks. The spatialization relative to the application of the regression tree on the city of Livorno (Fig. 3) highlights how the visual quality, estimated using the model of the regression tree, indicates high values in the area of the Terrazza Mascagni, in the two main squares of Livorno, Piazza Grande and Piazza della Repubblica, and in the area of the Fortress. The areas with a higher Visual Quality Index are mainly public spaces that show peculiar configurations of the site, the two main squares that are also wide bridges and the Terrazza Mascagni, a belvedere toward the sea.

3.2 From the Visual Quality Index to the rules for the urban design

The results of the spatialization of the VQI of the city of Livorno can be useful for identifying the rules for designing or regenerating urban spaces. The approach proposed by Alexander in his work "A pattern language" (Paterson and Connery, 1997) was chosen to analyse the characteristics of the areas with the highest visual quality, since the quality was assessed through a model that allowed to spatialise citizens' emotional perception. Alexander aimed to explore "the idea and principle of the process of designing by mass," and "to create a common language for designing and construction, in order to establish a process to let everybody participate in creating their own environment" (King, 1993: p.17). As Takashi *et al.* in the 1970s, Alexander (1979) observed that certain things were repeatedly seen in the form of architectures. He found that the relation among these things was a "pattern". The identification of the patterns was of fundamental support in the

Graph 3. Regression tree.
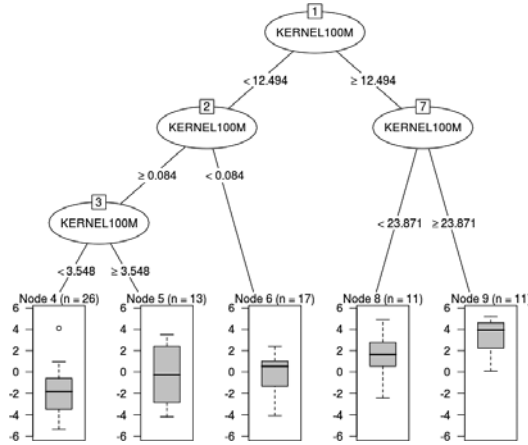
videos recorded during the administration of the questionnaire for the sampled points included in leaf 8 and 9 of the VQI map (figure 4). To select the patterns that most influenced the evaluation given by the respondents, the videos were analysed using the Graph-Based Visual Saliency (GBVS) method proposed by Itti and Koch (2001). It is based on the concepts of the neural mechanisms defined by Koch and Ullman (Itti, 2005; Itti, 2007; Itti and Koch, 2000; Koch and Ullman, 1985). In the perception of the cityscape, people freely observe the space around them without a specific objective (Dupont, Antrop and Vaneetvelde , 2015). As a consequence, the distribution of the localizations (patterns) that generate the most attention (on which the eye focuses) will be mainly guided by the content of the visual stimulus. To identify these areas of interest rapidly and systematically, a particularly interesting methodology recently proposed for landscape analysis is the so-called "saliency maps". The "saliency maps" can be described as computational focusing maps, which code for conspicuity or salience in a bottom-up mode in every position of an image. Salience is defined as the distinction of the perceptive quality with which an element in the world is distinguished from its neighbours and therefore immediately attracts attention (Itti, 2007). The GBVS method represents the most recent attempt to reconstruct the low-level instinctive process in the identification of the locations on which attention is concentrated into the geographical space depicted in an image. The GBVS method and the algorithms available for calculating the "saliency maps" were initially designed for the analysis of digital images. For this reason, 10 frames per minute were systematically extracted (one every 6 seconds) and the saliency map of each frame was calculated to adapt it to a video (Fig. 4).

The analysis of the high-salience areas of the video in the leaf areas 8 and 9 of Fig. 3 was firstly made by creating three-dimensional virtual diagrams drawn on the spherical photo. An example is available in the supplementary materials. This methodology allowed defining the 10 patterns shown in Fig. 5.

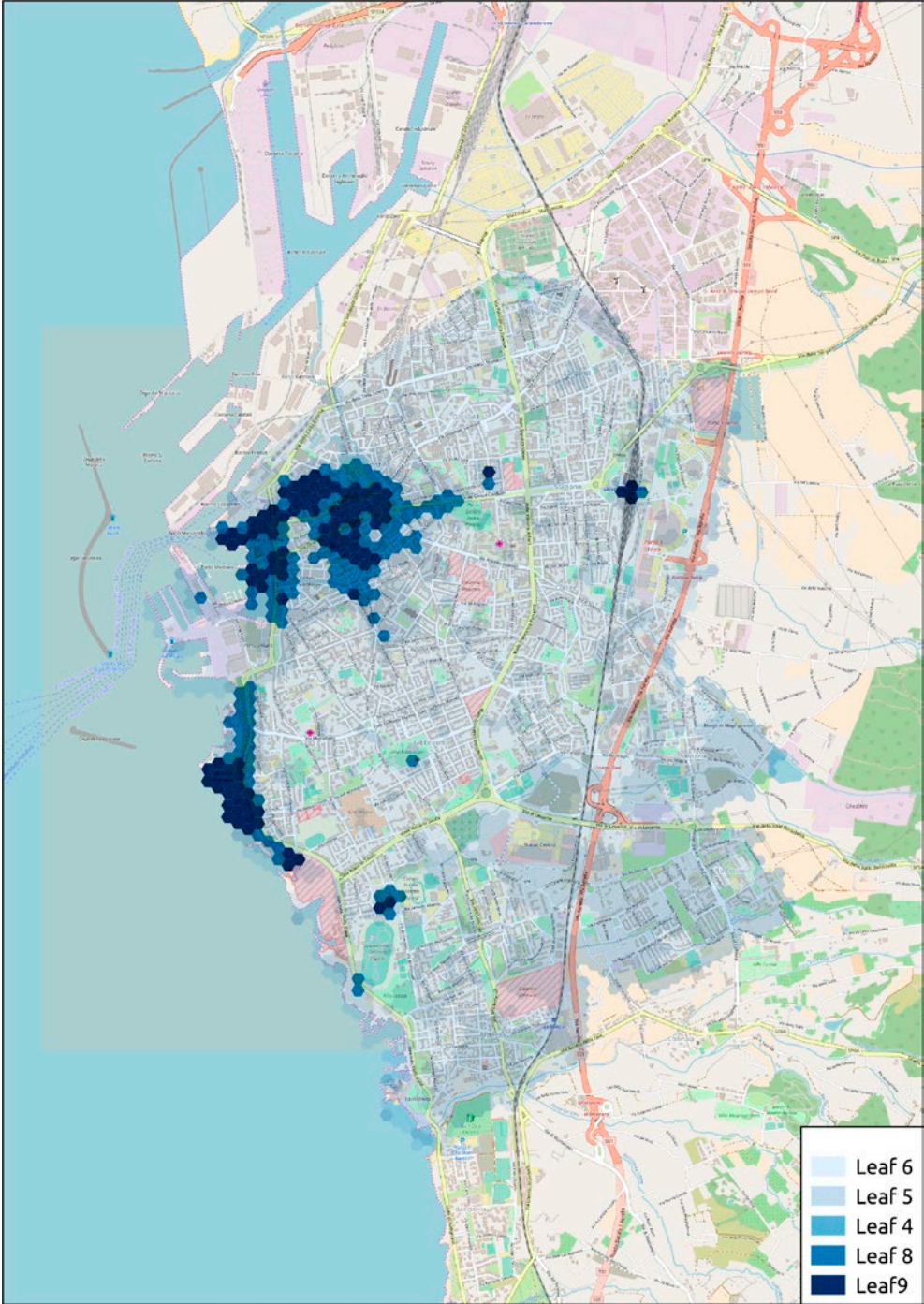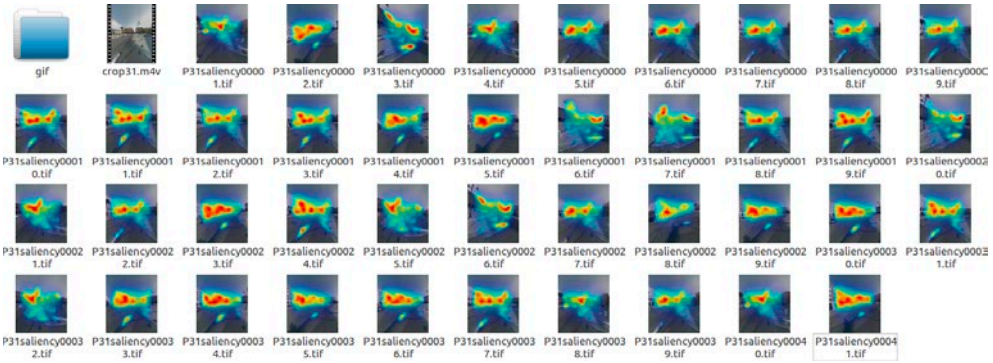Figure 3. Spatialization of the Visual Quality Index.

Figure 4. Video analysis using the GBVS method.

For each pattern the constituent elements and the dimensional ratios neces-
sary to reproduce it have been identified. In general, the survey showed that the
following elements are repeated in the areas with high VQI. A continuous street
wall characterizes the streets, and there is a boundary relationship between street
and water (canal or sea shore). Buildings are chromatically homogeneous, they
have decorative architectural ornaments (quoin, stringcourses, windows with ga-
bles, cornices, eaves, etc...) and are twice the width of the road. Often the streets
present elements of public interest much higher than the surrounding buildings
(e.g. churches, municipalities, etc.) at their beginning and/or at the end (vanishing
point). Road trees can be present (but not always).

A continuous street wall of buildings with height from 20 to 30 meters, often
with porticoes, characterizes the squares. The squares have a regular and sym-
metrical shape and present pedestrian spaces. The buildings have ornaments.
The squares with the highest VQI are the "bridge squares" typical of Livorno.
The "bridge square" is a paved square under which a canal flows. The presence
of the canal interrupts the street wall and opens wide perspective views. Quite
popular are also the small neighbourhood squares, characterized by half the size
of the main squares and by shops on the ground floor. Here too the street wall is
continuous, the shape is regular and symmetrical and there are ornaments on the
buildings too. The final phase of the work concerned the mapping of the patterns
detected in the city of Livorno (see Figure 5).

4. Discussion

The methodology used compares photo concentration with people's opinions
of three chosen sample. It allowed us to understand which characteristics of the
city are preferred by the community. The urban area of Livorno is divided into
three homogeneous areas according to the photo concentration's range and these
areas are confirmed by the questionnaire's results. This result defines the photo

Figure 5. Patterns and design rules in high visual quality urban space.
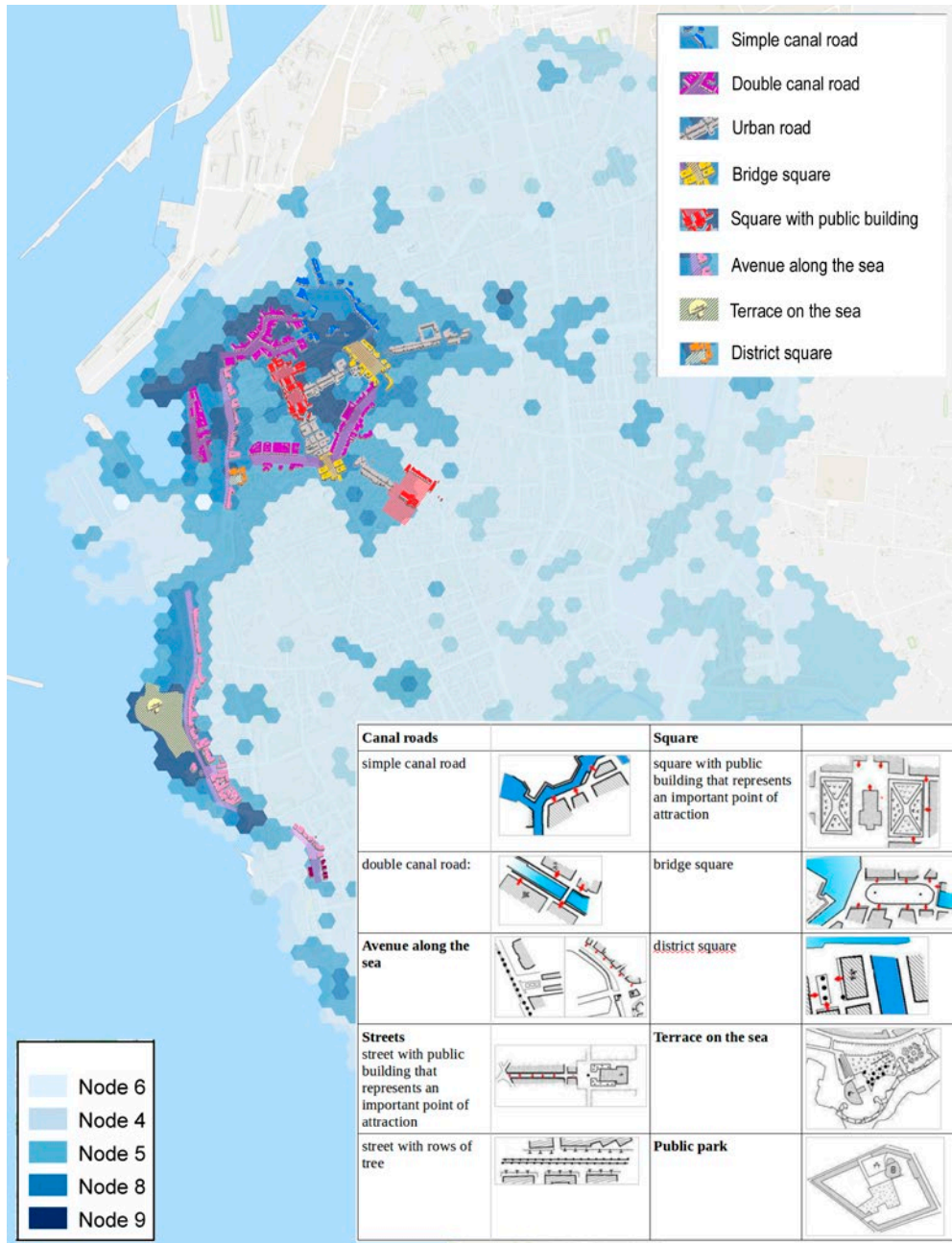
concentration as an index of urban quality. By analysing the locations taken as samples allows us to understand the characteristics of the city, their interaction and how they develop. Each location shows attributes that can bepositivelyor negatively perceived by the visitor. In any case, they represent important information to be used in urban planning to improve the overall quality of the city.

In particular, the results of this work can be used as a guideline in the process of creation and redevelopment of public open spaces in Livorno. The results demonstrate how subjective visual perception can help identifying and establishing the criteria for the design of high-quality public spaces. The quality patterns that have been identified suggest that functional, dynamic, associative, homogeneous, understandable, accessible, authentic spaces, with an emphasis on the identity elements, obtain visual appreciation on the users. On the other hand, spaces that do not have a defined visual identity in the structure of the city have been negatively evaluated. The research also highlighted how the use of open public spaces in Livorno as a primary visual resource of the city has potential, even though a large part of it needs revitalization, especially in terms of contents and dynamics.

The VQI map can be a useful tool for (Liu *et al.*, 2017):

- planning: municipal planners and administrators can achieve a more detailed and complete understanding of the urban environment. The map can provide good support for the definition of urban regeneration policies;
- research: researchers can further correlate the visual quality of the urban environment with socio-economic variables and the perception of quality of life;
- economic development: investors can detect the most suitable areas for the opening of restaurants and other public places. Local visitors can also benefit from these maps knowing more about high quality places to visit in the city and selecting more pleasant routes for their daily trips.

This method shows some positive and some negative aspects. The first ones are the low cost and the easy reproducibility of the methodology, while the second ones are the need to have a large number of photos available as inputs to make the methodology be effective and the fact that it cannot be implemented in those areas that are not crossed by the visitors.

Among the future objectives of this research we can certainly identify the following: to extend the methodology to unknown areas, to use the results to elaborate specific and precise urban policies and, lastly, to elaborate an automatic methodology to define specific indices able to describe the complexity of the cultural ecosystem services. This last element would allow us to compare the data with Flickr in order to have a ranking of the characteristics that most influence the aesthetic perception of a place.

5. Conclusions.

This research is based on the awareness of the loss of value of some urban spaces. These particular spaces do not favour the sense of belonging to the place

where people live. To address this issue, we need to solve the problem of assessing which configurations of space are the most suitable in a process of urban regeneration. The process of regeneration should be able to create emotions, which results in a more empathic space. In this respect, the research provides a method for evaluating the emotional qualities of urban configurations. Quality is sought in the interaction between the emotional aspects, activated by spatial perception, and the structural formal components, as they are displayed.

The first step in this study was to develop a photo density vector map. The map allowed the creation of a spatial sampling scheme. The map, obtained by spatially distributing the photos spontaneously posted on social networks by anonymous operators, determined an uneven density distribution. In particular, there was a high photo density on small areas and, conversely, a low photo density distributed in the remaining large part of the study area.

The analysis of the results of the given questionnaire, developed using Principal Components Analysis, allowed considerations on the concept of emotional content of specific urban configurations.

Through the PCA, which synthesizes the starting items of the semantic differential, two new significant variables have been defined for the evaluation of the quality of the spatial configurations of the urban environment: the formal and the emotional components. The emotional component, which has a very high percentage of variance, is significant in the description of the phenomenon, revealing that this empathetic relationship with the urban space is independent of the rational understanding of urban rules. Through the semantic differential items that saturate this component, it is possible to give a first definition of the concept of quality of spatial configurations. Space is positively perceived in an emotional way when there are formal components, with the following attributes: cared-for, wide, varied, special, open, decorated.

It was possible to interpret these formal components by studying those places that were found to be related to the emotional dimension.

Lastly, a Visual Quality Index has been developed to verify, in a more detailed way, whether and to what extent the new technologies, especially big data, can be useful indicators in defining the concept of quality of urban spatial configurations. By calculating the regression tree, it was possible to spatialise the relationship between the Visual Quality Index and the photo density.

Future developments in the research might investigate whether the information provided by big data on the quality of urban spatial configurations varies according to different contexts. The research represents a first step in that direction and could lead to more detailed and precise conclusions. First of all, it would be desirable to use the information coming from the crowdsourced geographic geodata through text mining procedures of the tags in their entirety. Currently the methodology considers only the geographical location of the photo. This would allow us to capture the unconscious aspects of the perception of the cityscape. On the other hand, the analysis of the photos taken with automatic methods of classification and categorizations of the images could allow studying the preferences expressed for particular urban patterns. Many authors have found that the visual

quality of the built environment is related to the social and economic characteristics of the population residing there. The co-evolution between the physical aspect and the social composition of the cities has for centuries aroused considerable interest on the side of scholars and supported most of the architectural and urban planning movements (Naik *et al.*, 2015). For example, Ordonez and Berg (2014) found a Pearson correlation coefficient of 0.51 between family income and judgement based on wealth environment and a coefficient of -0.36 between homicide statistics and safety judgement based on environment. Naik *et al.* (2015) found that the level of education in a neighbourhood is strongly related to urban quality. Another future development of this research could be to correlate the VQI with the levels of well-being. Lastly, the semantic differential and the evaluation methodologies of the VQI could be applied to virtual projects for the realization and regeneration of urban public spaces.

References

Abdi H. & Williams L. J. (2010). Principal component analysis. Wiley interdisciplinary reviews: computational statistics 2(4): 433-459.

Alexander C. (1979). The timeless way of building (Vol. 1). New York: Oxford University Press.

Alexander C., Ishikawa, S. & Silverstein M. (1977). A Pattern Language: Oxford University Press.

Arnheim R.(1969).Visual Thinking. Berkeley: University of California Press.

Asihara Y. (1983). The Aesthetic Townscape. Cambridge, MA: MIT Press.

Bailey T.C. & Gatrell A.C.(1995). Interactive Spatial Data Analysis. Longman Higher Education, Harlow.

Bonaiuto M., Fornara, F. & Bonnes, M. (2002). Indexes of perceived residential environment quality and neighborhood attachment in urban environments: a confirmation study on the city of Rome Landscape and Urban Planning 988: 1-12.

Borruso, G. (2013). Cartografia e informazione geografica "2.0 e oltre", webmapping, webgis. un'introduzione. EUT Edizioni Università di Trieste.

Brown G., de Bie K. & Weber D. (2015). Identifying public land stakeholder perspectives for implementing place-based land management. Landscape and Urban Planning 139: 1-15.

Bruns A. (2006). Towards produsage: Futures for user-led content production. Canadian Geographer 47(2): 185-200.

Carmona M., Heath, T., Oc T. & Tiesdell S. (2012). Public Places - Urban Spaces: The Dimensions of Urban Design. Oxford: Architectural Press.

Cauchy, A. (1829). Sur l'équation à l'aide de laquelle on détermine les inégalités séculaires des mouvements des planets. Exerc. Math. 4: 140-160.

Chen J. & Shaw S.L. (2016) Representing the Spatial Extent of Places Based on Flickr Photos with a Representativeness-Weighted Kernel Density Estimation. In: Miller J., O'Sullivan D., Wiegand N. (eds) Geographic Information Science. GIScience 2016. Lecture Notes in Computer Science, vol 9927. Springer, Cham.

Chen J. & Takamatsu M. (2016). The Evaluation of Urban Landscape upon Japanese Representative LRT Cities Using Visual Engineering. American Journal of Engineering Research 5(2): 136-144.

Collier Jr J. (1967). Visual Anthropology: Photography as a Research Method. New York, Chicago, San Francisco. Atlanta, Dallas, Montreal, Toronto, London: Holt, Rinehart and Winston, Inc.

Dakin S. (2003). There's more to landscape than meets the eye: towards inclusive landscape assessment in resource and environmental management. The Canadian Geographer/Le Géographecanadien 47(2): 185-200.

Danese M., Lazzari, M. & Murgante B. (2008, June). Kernel density estimation methods for a geo-

statistical approach in seismic risk analysis: The case study of Potenza hilltop town (Southern italy). In International Conference on Computational Science and Its Applications (pp. 415-429). Springer, Berlin, Heidelberg.

de Val G. D. L. F., Atauri J. A., & de Lucio, J. V. (2006). Relationship between landscape visual attributes and spatial pattern indices: a test study in Mediterranean-climate landscapes. Landscape and urban planning 77(4): 393-407.

Di Franco G. & Marradi A. (2003). Analisi fattoriale e analisi in componenti principali. Bonanno.

Dunkel A. (2015). Visualizing the perceived environment using crowdsourced photo geodata. Landscape and urban planning 142: 173-186.

Dupont L., Antrop M. & Van Eetvelde, V. (2015). Does landscape related expertise influence the visual perception of landscape photographs? Implications for participatory landscape planning and management. Landscape and Urban Planning 141: 68-77.

Feick R. & Robertson C. (2015). A multi-scale approach to exploring urban places in geotagged photographs. Computers, Environment and Urban Systems 53: 96-109.

Fitzgerald F. S. (2001). Tender is the Night. Penguin UK.

Franz G, von der Heyde M & Bülthoff HH, (2005).Predicting experiential qualities of architecture by its spatial properties. In Designing Social Innovation: Planning, Building, Evaluating Eds Martens, B, Keul, AG, (Hogrefe and Huber, Cambridge, MA) pp. 157-166.

Geddes P. (1971). Cities in evolution: an introduction to the town planning movement and to the study of civics. Harper & Row.

Gibson J. J. (1979). The Ecological Approach to Visual Perception.Boston :Houghton Mifflin.

Green R. (1999). Meaning and form in community perception of town character. Journal of Environmental Psychology 19(4): 311-329.

Hauthal E. & Burghardt D. (2016). Using VGI for analyzing active ties and emotions of locals and tourists. In Link-VGI workshop in connection with the AGILE.

Hsu S.H, Chuang M.C., Chang C.C. (2000). A semantic differential study of designers' and users' product form perception International Journal of Industrial Ergonomics 25: 375-391.

Hung T. F. & Nieh C. K. (2009). Application of Semantic differential technique to evaluate Kansei image in architectural design. In Proceedings of the international design research conference–IASDR

Hotelling H. (1933). Analysis of a complex of statistical variablesinto principal components. J Educ Psychol 1933, 25:417–441.

Itti L. (2005). Models of bottom-up attention and saliency. In Neurobiology of attention (pp. 576-582).

Itti L. (2007). Visual salience. Scholarpedia 2(9): 3327.

Itti L. & Koch C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. Vision Research 40: 1489–1506.

Itti L. & Koch C. (2001). Computational modelling of visual attention.Nature Reviews Neuroscience 2: 194-203.

Imamoglu Ç. (2000). Complexity, liking and familiarity: architecture and non-architecture Turkish students' assessments of traditional and modern house facades. Journal of Environmental Psychology 20(1): 5-16.

Jin X., Gallagher A., Cao L., Luo J. & Han J. (2010, October). The wisdom of social multimedia: using flickr for prediction and forecast. In Proceedings of the 18th ACM international conference on Multimedia (pp. 1235-1244). ACM.

Jordan, C. (1874). Mémoire sur les formes bilinéaires. Journal de mathématiques pures et appliquées 19: 35-54.

Kang J. & Zhang M. (2010). Semantic differential analysis of the soundscape in urban open public spaces.Building and environment, 45(1): 150-157.

King I. F. (1993).Christopher Alexander and contemporary architecture.A+ U Pub. Co.

Koch C., Ullman S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry Human Neurobiology 4: 219-227.

Kuliga S. F., Thrash T., Dalton R. C. & Hoelscher C. (2015). Virtual reality as an empirical research tool—Exploring user experience in a real building and a corresponding virtual model. Computers, Environment and Urban Systems 54: 363-375.

Lim E. M., Honjo T., & Umeki K. (2006).The validity of VRML images as a stimulus for landscape assessment. Landscape and Urban Planning 77(1-2): 80-93.

Liu L., Silva E. A., Wu C.& Wang H. (2017). A machine learning-based method for the large-scale evaluation of the qualities of the urban environment. Computers, Environment and Urban Systems 65: 113-125.

Lothian A. (1999). Landscape and the philosophy of aesthetics: is landscape quality inherent in the landscape or in the eye of the beholder?.Landscape and urban planning 44(4): 177-198.

Lovett A., Appleton K., Warren-Kretzschmar B. & Von Haaren C. (2015). Using 3D visualization methods in landscape planning: An evaluation of options and practical issues. Landscape and Urban Planning 142: 85-94.

Lynch K. (1960). The image of the city (Vol. 11). MIT press.

Maffei L., Masullo M., Pascale A., Ruggiero G. & Puyana Romero V."(2015). On the Validity of Immersive Virtual Reality as tool for multisensory evaluation of urban spaces. Energy Procedia 78: 471-476.

Mallgrave H. F. (2015). L'empatia degli spazi. Architettura e neuroscienze, Raffaello Cortina, Milano.

Mumford L. (2017). Storia dell'utopia, Feltrinelli, Milano.

Naik N., Kominers S. D., Raskar R., Glaeser E. L. & Hidalgo C. A. (2015). Do people shape cities, or do cities shape people? The co-evolution of physical, social, and economic change in five major US cities (No. w21620). National Bureau of Economic Research.

Nov O., Naaman M. & Ye C. (2010). Analysis of participation in an online photo sharing community: A multidimensional perspective. Journal of the Association for Information Science and Technology 61(3): 555-566.

Ordonez V. & Berg T. L. (2014). Learning high-level judgments of urban perception. In European Conference on Computer Vision (pp. 494-510). Springer, Cham.

Orenstein D. E., Zimroni H. & Eizenberg E. (2015). The immersive visualization theater: A new tool for ecosystem assessment and landscape planning. Computers, Environment and Urban Systems 54: 347-355.

Orland B., Budthimedhee K. & Uusitalo J. (2001). Considering virtual worlds as representations of landscape realities and as tools for landscape planning. Landscape and urban planning 54(1): 139-148.

Osgood C. E., Suci G. J. & Tannenbaum P. H. (1957). The Measurement of Meaning. Urbana: University of Illinois Press.

Pallasmaa J., Mallgrave H. F., Robinson S. & Gallese V. (2015). Architecture and Empathy.

Paterson D. & Connery K. (1997). Reconfiguring the edge city: The use of ecological design parameters in defining the form of community. Landscape and Urban Planning 36(4): 327-346.

Patil G. P., Myers W. L., Luo Z., Johnson G. D. & Taillie C. (2000). Multiscale assessment of landscapes and watersheds with synoptic multivariate spatial data in environmental and ecological statistics. Mathematical and Computer Modelling, 32(1-2): 257-272.

Pearson K. (1901). On lines and planes of closest fit to systems of points in space. Philos Mag A 6: 559-572

Perovic, S. & Folic N. K. (2012).Visual perception of public open spaces in Niksic.Procedia-Social and Behavioral Sciences 68: 921-933.

Piga B.E.A., Tangredi C., Signorelli V. (2011b). Investigating the traveler experience: the use of eye-tracking for urban design purposes, Proceeding of the 10th Conference of the European Architectural Envisioning Association, Delft, Netherlands, 14-17/09/2011, Delft: Delft Digital Press: 161-168.

Piga B. E., Cibien L., Legnani A., Mungo S., Zappa D., Signorelli V. & Morello E. (2013).Asense-archive of urbanspace.In EAEA-11 conference (pp. 251-258).

Saragosa C. (2016) Il sentiero di Biopoli. L'empatia nella generazione della città, Donzelli, Roma.

Scott M. J. & Canter D. V. (1997). Picture or place? A multiple sorting study of landscape.Journal of environmental psychology 17(4): 263-281.

Settis S. (2017). Architettura e democrazia. Paesaggio, città, diritti civili, Einaudi, Torino.

Spyrou E. & Mylonas P. (2016). A survey on Flickr multimedia research challenges. Engineering Applications of Artificial Intelligence 51: 71-91.

Sontag, S. (1977). On photography. New York: Farrar, Straus and Giroux.

Stewart GW. (1993). On the early history of the singular value decomposition. SIAM Rev. 35: 551-566.

Tempesta T. (2009). Economia del paesaggiorurale. La valutazione del danno ambientale e paesaggistico, Villa Umbra-Regione Umbria: 9-10.

Tenkanen H., Di Minin E., Heikinheimo V., Hausmann A., Herbst M., Kajala L. & Toivonen T. (2017). Instagram, Flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected areas. Scientific reports 7(1): 17615.

Torgo L. (2016). Data mining with R: learning with case studies. CRC press.

White D., Kimerling J. A. & Overton S. W. (1992). Cartographic and geometric components of a global sampling design for environmental monitoring. Cartography and geographic information systems 19(1): 5-22.

Zhou X., Xu C. & Kimmons B. (2015). Detecting tourism destinations using scalable geospatial analysis based on cloud computing platform. Computers, Environment and Urban Systems 54: 144-153.

Zube E. H., Sell J. L. & Taylor J. G. (1982). Landscape perception: research, application and theory. Landscape planning 9(1): 1-33.