

Bag of Visual-words と SVM を用いたキュビズム絵画の分類*

三 角 真**
 折 居 英 章***
 Tanjina SHARMIN****
 三 島 健 司****
 鶴 岡 智 昭**

Computer Aided Recognition of Cubism Painting

Makoto MISUMI**, Hideaki ORII***, Tanjina SHARMIN****, Kenji MISHIMA****
 and Tomoaki TSURUOKA**

Abstract

In recent years, it has been required to record the traditionally important painting as digital data. Classification of the painting style of painting by a computer is required. The painting style of painting is characterized by such as colors, brushstroke and texture of the painting. When we classify the huge data of the painting, we should recognize the painting style of painting by using an image recognition method. In this study, we discuss a new approach to achieve the objective classification and identification of the painting style of paintings using the image recognition approach. "Bag of Visual-words" was used in this work. That has been widely used in the field of image recognition. At the learning of the feature, Support Vector Machine (SVM) was used as the machine learning method. In the experiment, accuracy of the classification attains 72%.

Key Words : Cubism, Bag of Visual-words, Support Vector Machine, Painting

1. 序論

絵画をデジタルデータとして記録することが近年求められている。その際、絵画の画風を自動的に分類することが必要不可欠である。絵画の画風は、筆の運び、色使いなどで特徴づけられ、絵に表現された画家個人又は流派の特色や傾向を表しているため芸術性を測る上で大きな要素の一つとなっている。画風について客観的な数値指標を用いて識別・解析を行うことは、芸術性を考察する上で非常に重要である。一方、画風は絵画作品上で色の配置や模様が複雑に構成されたものとして表現され

るため、数値では表現しづらいという現状もある。

そこで本研究では、画像認識的なアプローチを用い、絵画作品の画風の客観的な分類・識別を実現する新たな手法について検討する。本手法では、近年画像認識の分野で広く用いられている「Bag of Visual-words」の枠組みに基づき絵画画像から特徴抽出を行い、Support Vector Machine (SVM) を用いた機械学習によりその特徴の学習を行う。実験では、Web から収集した典型的なキュビズム絵画データと、アールヌーボーや表現主義派などのキュビズム絵画と画風の異なる絵画画像の分類実験を行い、提案手法の有効性を検証した。

2. 提案手法

本研究では、画像認識の分野で広く用いられている「Bag of Visual-words」の枠組みにもとづき、一般絵画の分類を行った¹⁻³⁾。Bag of Visual-words (BoVW) では、

* 平成 27 年 10 月 31 日受付
 ** 電子情報工学科
 *** 電気工学科
 **** 化学システム工学科

画像内に頻出する画像パターン (Visual-Words) を求める「特微量辞書の作成」処理, 識別対象の種類 (クラス) の特徴を統計的に学習する「学習」処理, 入力画像の所属クラスを推定する「分類」処理を組み合わせることで画像の

分類を行う。図1に手法の概念図を, 図2に処理の流れを示す。

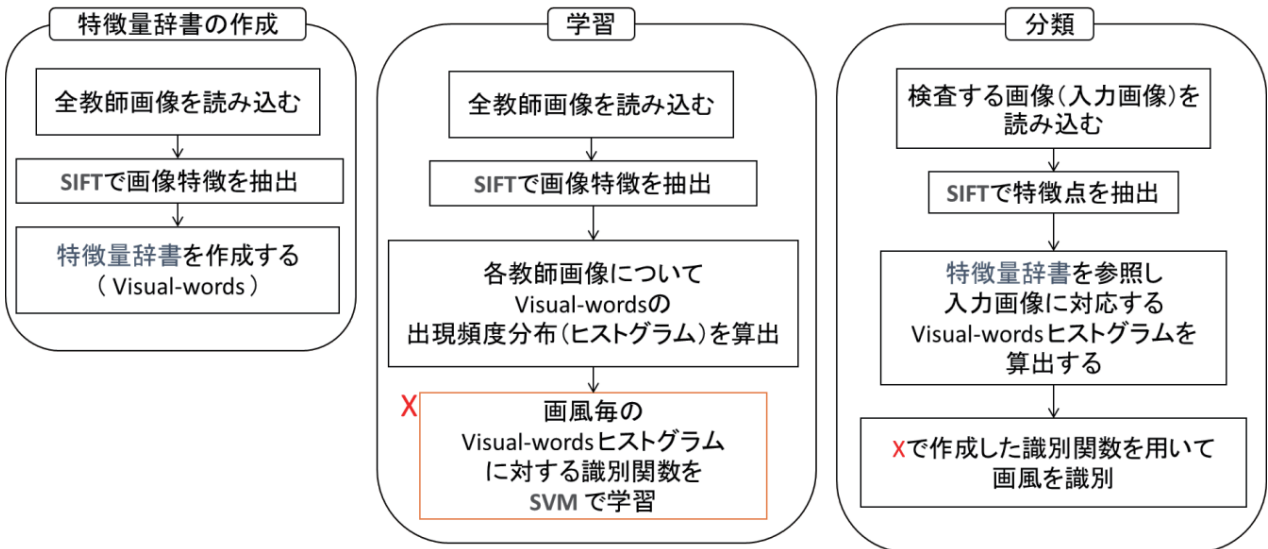


図1 提案手法全体の流れ

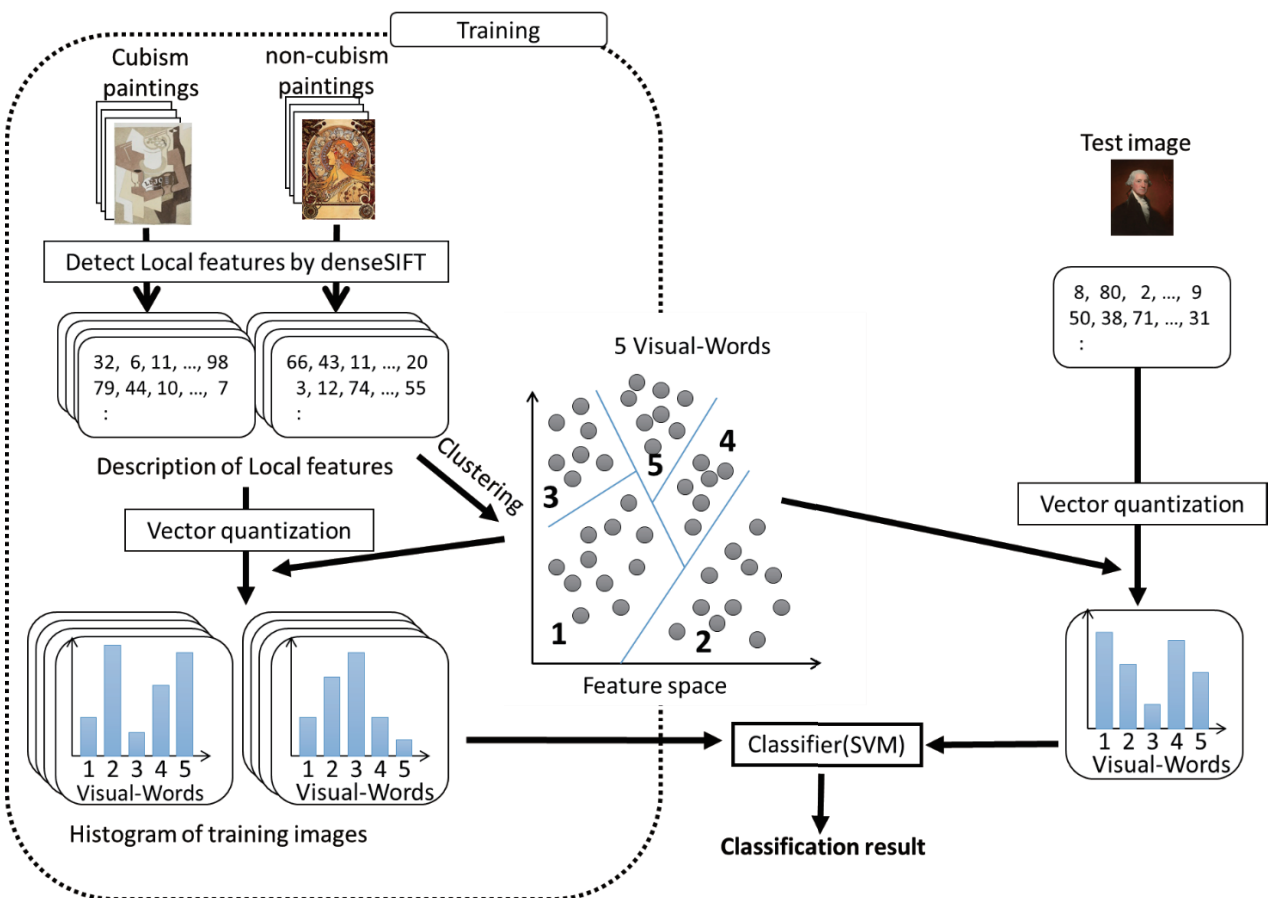


図2 提案手法の概念図

2.1. 局所画像特徴量

本研究では、局所画像パターンの特徴記述に用いる画像特徴量として Scale-Invariant Feature Transform (SIFT)^{2,8,10,11)} 特徴量を用いた。SIFT 特徴量は画像の輝度勾配を用いて記述される特徴量で、画像パターンの回転変化や明るさ変化に頑健な特徴記述が可能という特長を持つ。SIFT では、図 3 に示すように、画像上の任意の点をキーポイントとして、キーポイントを中心とした任意サイズの矩形領域を選択し、矩形領域を 16 個の小さな矩形領域に分割して、各矩形領域において勾配強度を重みとした 8 方向の重み付きヒストグラムを算出しそれを連結したものを局所特徴量とする。その際、キーポイントを中心とした領域のオリエンテーション (方向) を領域全体の勾配強度分布から求め、特徴記述の際にオリエンテーションにより向き正規化を行うことで回転に不変となる。以下にその詳細について説明する。

キーポイントのオリエンテーションを求めるには、まず画像 $I(u,v)$ の勾配強度 $m(u,v)$ と勾配方向 $\theta(u,v)$ を以下の式により求める。

$$m(u,v) = \sqrt{f_u(u,v)^2 + f_v(u,v)^2} \quad (1)$$

$$\theta(u,v) = \tan^{-1} \frac{f_u(u,v)}{f_v(u,v)} \quad (2)$$

$$\begin{cases} f_u(u,v) = I(u+1,v) - I(u-1,v) \\ f_v(u,v) = I(u,v+1) - I(u,v-1) \end{cases} \quad (3)$$

そして、キーポイント周辺領域における勾配強度 $m(x,y)$ と勾配方向 $\theta(x,y)$ から勾配強度に関する重み付き方向ヒストグラム h を以下の式により作成する。

$$h_{\theta'} = \sum_x \sum_y w(x,y) \cdot \delta[\theta', \theta(x,y)] \quad (4)$$

$$w(x,y) = G(x,y,\sigma) \cdot m(x,y) \quad (5)$$

ここで、 h_{θ} は全方向を 36 方向に量子化したヒストグラムである。 $w(x,y)$ はある局所領域の画素 (x,y) での重みであり、キーポイント周辺領域のサイズと同等のスケ

ールのガウス窓 $G(x,y,\sigma)$ と勾配強度 $m(x,y)$ から求める。 $\delta[\cdot, \cdot]$ は Kronecker のデルタ関数で、勾配方向 $\theta(x,y)$ が量子化した方向 θ' に含まれるとき 1 を返す。この 36 方向のヒストグラムの最大値から 80% 以上となるピークをキーポイントのオリエンテーションとして割り当てる。ガウス窓による重み付けにより、キーポイントに近い画素が持つ勾配強度がオリエンテーションの決定に対して強く反映される。

検出したオリエンテーションを基に、SIFT 特徴量を記述する。まず、キーポイント周辺領域画像をキーポイントのオリエンテーション方向に回転する。特徴量の記述には、キーポイント周辺画素が持つ勾配情報を用いる。キーポイント周辺領域を一辺が 4 ブロックの計 16 ブロックに分割し、各ブロック 8 方向 (45 度ずつ) の勾配方向ヒストグラムを作成する。この手順はキーポイントのオリエンテーション算出と同様に行う。したがって、 $4 \times 4 = 16$ ブロックの各領域で 8 方向の勾配強度に関する重み付き方向ヒストグラムを作成するため、特徴量としては $4 \times 4 \times 8 = 128$ 次元の特徴ベクトルが求まる。この特徴ベクトルはベクトルの総和で正規化する。このように、オリエンテーションによる方向の正規化、輝度勾配を用いた特徴記述を用いるため、SIFT 特徴量は回転不変性・照明変化頑健性を持つ。

一般的に、SIFT ではキーポイントの検出に Difference-of-Gaussian(DoG) 処理を用いる。これは視覚的に顕著な特徴を抽出し特徴量に拡大・縮小の不変性を持たせるためであるが、その場合 BoVW を実現するのに十分な特徴点の数が得られないことがある。それに対して、画像をグリッドに区切りそれらの中心をキーポイントとし画像から密に特徴量を得る dense sampling という方式がある。本方式では 1 枚の画像全体から均一に多数の局所画像特徴量を得ることができるため、本研究ではこの dense sampling を用いる dense SIFT^{2,8,9)} を画像からの特徴量抽出手法として採用した。

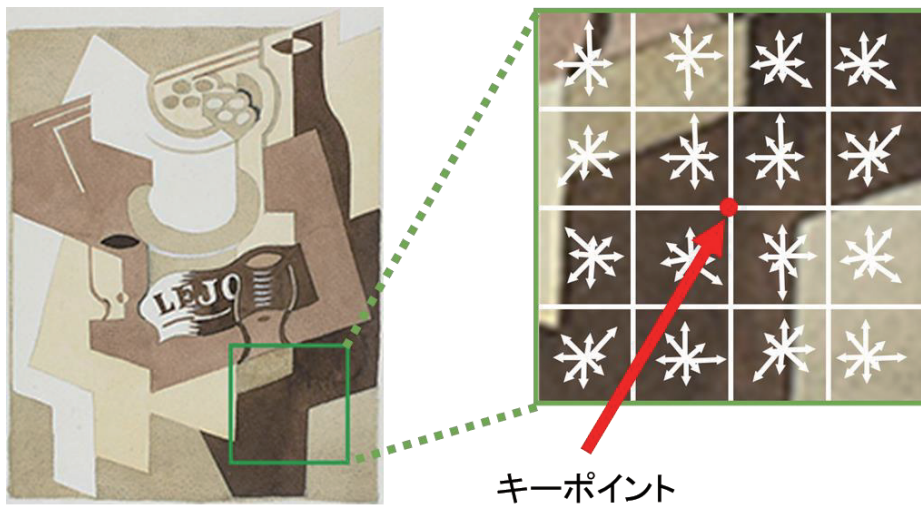


図 3 SIFT 特徴量 (代表的キュビズム絵画: 12))

2.2. Visual-Words 辞書の作成

BoVW では、まず、多数の一般画像を用いて画像に頻出する局所画像パターンを求め、Visual-Words 辞書とする。本研究では、無作為に選んだ多数の絵画画像から dense SIFT により特徴量を抽出し、その全特徴量に対して k-means アルゴリズムを用いてクラスタリングを行うことにより得られるクラスタ中心特徴量を Visual-Words 辞書とした。

2.3. 対象画像のヒストグラム算出

前節で得られた Visual-Words 辞書を基に、任意画像から dense SIFT で得られる画像の各局所特徴量について最も近い visual-word を求めれば、その出現頻度分布をヒストグラムで表現することができる。BoVW ではこれを全体画像特徴量として用いる。したがって、本研究でも各絵画の画像データから dense SIFT に基づいたコードパターンの出現頻度によって表されるヒストグラムを算出し、絵画画像の全体画像特徴量とした。各種類の絵画の画像データから得られたコードパターンの出現頻度のヒストグラムの一例を図4に示す。図4(a)は、図3に示したキュビズム絵画から算出したもので、各局所特徴に対してコードパターンの出現頻度によって表されるヒストグラムの偏りが大きく、単純な模様からなるキュビズム絵画の特徴をよく表現している。また、図4(b)は、複雑なアールヌーボー絵画から算出したコードパターンの出現頻度のヒストグラムであり、多くの局所特徴を有しており、そのヒストグラムの分布も偏りがすくなく多様なものとなっている。

2.4. Support Vector Machine を用いた分類・識別

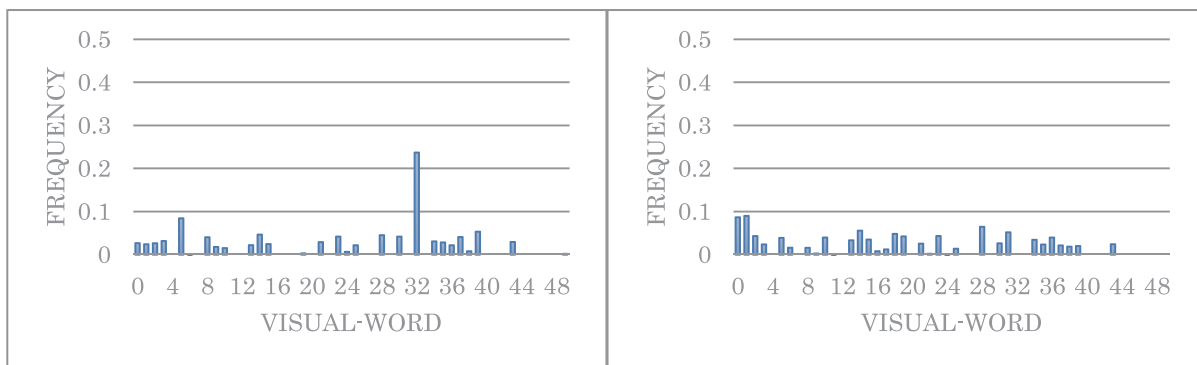
前節で述べたように、各絵画の Visual Words ヒストグラムはその画風を反映した分布となる。しかしながら、その分布の特徴を定量的に定義することは難しい。そこで本研究では、画風を表す Visual Words ヒストグラムの特徴を統計的学習処理により求める。Support Vector Machine (SVM) ⁴⁷⁾ は2クラス分類を行う機械学習の手法である。カーネル法を用いてデータ空間を高次特徴空間に写像することで線形問題だけでなく非線形問題にも適用できるのが特長で、マージン最大化という考えに基づき識別面が各クラスの境界付近に存在するデータと最大距離を持つように学習が行われる。SVM を用いたデータの識別は、次式のような識別関数を用いて行う。

$$d(\mathbf{x}) = \Phi(\mathbf{m}, \mathbf{x}) + b \tag{6}$$

ここで、 \mathbf{x} は入力データ、 \mathbf{m} は Support Vector、 $\Phi(\)$ はベクトル \mathbf{m} と \mathbf{x} の高次特徴空間での内積値を算出するカーネル関数、 b はバイアス項である。一般的に、 $d(\mathbf{x})$ の値が正負どちらの値をとるかでデータ \mathbf{x} のクラスを識別する。高次特徴空間を定義するカーネル関数にはいくつか種類があるが、本研究では以下の RBF (Radial Basis Function) kernel を用いた。

$$\Phi_{\text{RBF}}(\mathbf{m}, \mathbf{x}) = \exp(-\gamma \cdot |\mathbf{m} - \mathbf{x}|^2) \tag{7}$$

SVM を用いて画風の分類を行うため、まず、「キュビズム絵画」と「それ以外」の2つの学習用教師画像群を用意し、それぞれの画像について Visual Words ヒストグラム特徴算出し学習データとした。この学習により、キュ



(a) キュビズム絵画より算出

(b) アールヌーボー絵画より算出

図4 各種絵画から算出し局所特徴コードパターンの出現頻度によって表されるヒストグラム例

ビズム絵画の Visual Words ヒストグラム分布の特徴が識別関数として抽出される。そして、画像認識時にはこの識別関数を用いて分類を行うことが可能となる。

3. 実験

我々は、キュビズム絵画のデータとして en.wikipedia.org の Cubism に掲載された 18 枚、アールヌーボー絵画のデータとして 11 枚、表現主義派の絵画として 13 枚収集した。それら全ての画像について横幅を 200 ピクセルにリサイズし、1 ピクセル間隔でキーポイントを設定し、そのキーポイントを中心とする 32 ピクセル、45 ピクセル、64 ピクセル四方の矩形領域から denseSIFT を用いて特徴量を抽出し、k-means を用いて 50 語、100 語、500 語の Visual-word 辞書を生成した。

収集した絵画データをランダムに 2 グループに分け、RBF カーネルを用いた SVM で Visual-Word 辞書ごとに交差検定を行った。

4. 実験結果と考察

図 5 に、SVM を用いた Visual-Word の辞書サイズ別の分類結果を示す。この結果から、本実験においては 100 個の Visual-Word の辞書を用いることで 70% 以上の正解率でキュビズム画像と、非キュビズム画像を分類可能であることが分かる。

5. おわりに

本研究では、絵画の画風の客観的な分類・識別を目的として、近年画像認識の分野で広く用いられている「Bag of Visual-words」枠組みに基づいて絵画画像から特徴量抽出を行い、教師有り機械学習である SVM を用いて学習と分類を行った。その結果、Web から収集した典型的なキュビズム絵画と、アールヌーボーや現実主義派などのキュビズム絵画と画風の異なる絵画画像の分類において、少量の絵画データの分類であっても 70% 以上の正解率を得ることができた。

謝辞

この研究の一部は、公益財団法人 日本食品化学研究振興財団の研究助成金 (2015 年度) ならびに、文部科学省の科研費基盤 (C)(研究課題番号 26420770) の補助を受けた。

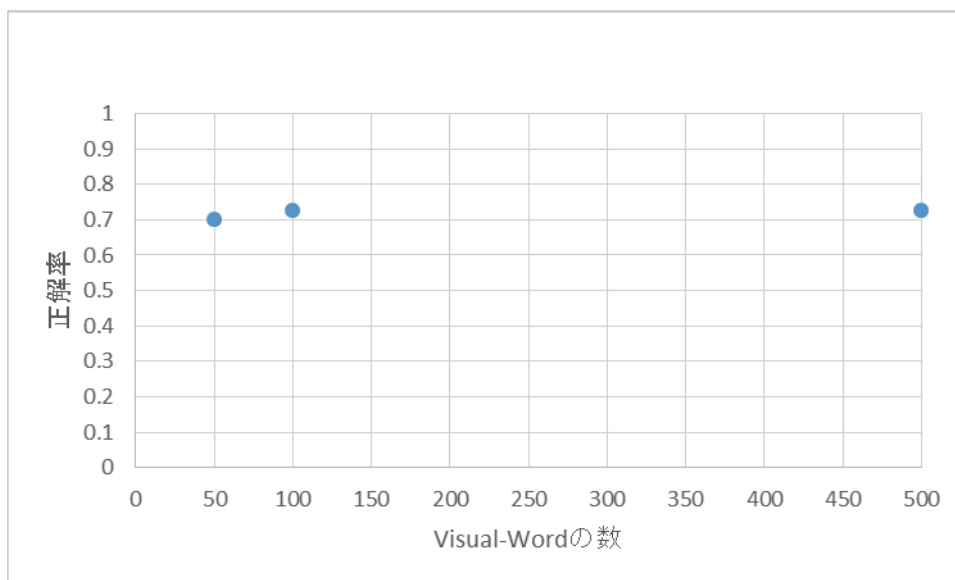


図 5 キュビズム画像交差検定の結果

参考文献

- 1) Csurka, G., Dance, C.R., Fan, L., Willamowski, J., Bray, C.: “Visual categorization with bags of keypoints”, European Conference on Computer Vision (ECCV2004) Workshop on Statistical Learning in Computer Vision, pp. 59–74, 2004.
- 2) Nowak, E., Jurie, F., Triggs, B.: “Sampling strategies for bag-of-features image classification.”, Computer Vision ECCV-2006, Lecture Notes in Computer Science, vol. 3954. Springer, Berlin/Heidelberg, pp. 490–503, 2006.
- 3) Lazebnik, S., Schmid, C., Ponce, J.: “Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories.”, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 2169–2178, 2006.
- 4) Vapnik, V.N: “Statistical Learning Theory.”, Wiley, New York, 1998.
- 5) John Shawe-Taylor, N.C.: “An Introduction to Support Vector Machines: And Other Kernel-based Learning Methods.”, Cambridge University Press, Cambridge, New York, 2000.
- 6) Schölkopf, B., Smola, A.J.: “Learning with kernels: support vector machines, regularization, optimization, and beyond.”, MIT Press, Cambridge, Mass, 2002.
- 7) Steinwart, I., Christmann, A.: “Support Vector Machines”, Springer, New York, 2008.
- 8) Fei-Fei, L., Perona, P.: “A bayesian hierarchical model for learning natural scene categories.”, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005, pp. 524–531, 2005.
- 9) Bosch, A., Zisserman, A., Muoz, X.: “Image classification using random forests and ferns.”, IEEE 11th International Conference on Computer Vision, 2007. ICCV 2007, pp. 1–8, 2007.
- 10) Jurie, F., Triggs, B.: “Creating efficient codebooks for visual recognition.”, Tenth IEEE International Conference on Computer Vision, 2005. ICCV 2005, pp. 604–610, 2005.
- 11) Herve, N., Boujemaa, N., Houle, M.E.: “Document description: what works for images should also work for text?”, Multimedia Content Access: Algorithms and Systems III, SPIE, pp. 72550B–72550B-12, 2009.
- 12) Juan Gris: “Still Life”, Art Institute of Chicago, 1919.
Emanuel M Ologeanu: “El Vagabundo – cubist painting”.