BESCHWERDEKAMMERN
DES EUROPÄISCHEN
PATENTAMTS

BOARDS OF APPEAL OF
THE EUROPEAN PATENT
OFFICE

CHAMBRES DE RECOURS
DE L'OFFICE EUROPÉEN
DES BREVETS

**Internal distribution code:**

(A) [ - ] Publication in OJ
(B) [ - ] To Chairmen and Members
(C) [ - ] To Chairmen
(D) [ X ] No distribution

**Datasheet for the decision
of 20 March 2023**

**Case Number:**              T 0814/20 - 3.5.06

**Application Number:**       15195657.0

**Publication Number:**       3035239

**IPC:**                      G06K9/00

**Language of the proceedings:**   EN

**Title of invention:**
ADAPTED VOCABULARIES FOR MATCHING IMAGE SIGNATURES WITH FISHER
VECTORS

**Applicant:**
Conduent Business Services, LLC

**Headword:**
Adapted Visual Vocabularies/CONDUENT

**Relevant legal provisions:**
EPC Art. 56, 84, 123(2)

**Keyword:**
Inventive step - technical purpose - effect made credible
within the whole scope of claim (yes)

**Decisions cited:**
G 0001/19

**Catchword:**

Beschwerdekammern

Boards of Appeal

Chambres de recours

Boards of Appeal of the
European Patent Office
Richard-Reitzner-Allee 8
85540 Haar
GERMANY
Tel. +49 (0)89 2399-0
Fax +49 (0)89 2399-4465

Case Number: **T 0814/20 - 3.5.06**


D E C I S I O N
of Technical Board of Appeal 3.5.06
of 20 March 2023


**Appellant:**                   Conduent Business Services, LLC
(Applicant)                      2828 North Haskell Avenue
                                 Building 1, Floor 9
                                 Dallas, TX 75204 (US)


**Representative:**              Gill Jennings & Every LLP
                                 The Broadgate Tower
                                 20 Primrose Street
                                 London EC2A 2ES (GB)


**Decision under appeal:**       **Decision of the Examining Division of the
                                 European Patent Office posted on 20 September
                                 2019 refusing European patent application No.
                                 15195657.0 pursuant to Article 97(2) EPC.**


**Composition of the Board:**

**Chairman**        M. Müller
**Members:**        T. Alecu
                    B. Müller

## Summary of Facts and Submissions

I.      The appeal is against the decision of the Examining
        Division to refuse the application. With the grounds of
        appeal the Appellant requested that the decision of the
        Examining Division be set aside and that a patent be
        granted on the basis of a main request, which was
        identical to the main request underlying the appealed
        decision, or on the basis of one of five auxiliary
        requests.

II.     In its decision, the Examining Division denied novelty
        or inventive step of all requests in comparison with
        document

        D3: EP 2 065 813 A1, also in view of
        D4: SANCHEZ JORGE ET AL: "Image Classification with the
            Fisher Vector: Theory and Practice".

        It further found a lack of support (Article 84 EPC) and
        an extension beyond the content of the application as
        originally filed (Article 123(2) EPC) for the first
        auxiliary request underlying its decision.

III.    In a communication accompanying a summons to oral
        proceedings the Board provided its provisional opinion
        that, for all requests, an inventive step could not be
        acknowledged for the subject matter of claim 1 because
        it did not provide a technical contribution to the art,
        Article 56 EPC. The Board also considered that the main
        and the first three auxiliary requests lacked support,
        Article 84 EPC, and that the second to fifth auxiliary
        requests did not comply with Article 123(2) EPC.

IV.     With its reply of 13 February 2023, the Appellant
        submitted several auxiliary requests.

V.      The Board indicated with its communications of
        1 March 2023 and 6 March 2023 that the application
        could be allowed to proceed to grant on the basis of a
        request labeled C4, but that further amendments to the
        claims and the description were necessary for the
        application to be compliant with the requirements of
        Articles 84 (clarity, conciseness and determination of
        the protection sought) and 123(2) EPC.

VI.     In response, the Appellant provided a first version of
        an amended set of claims on 3 March 2023 and a full re-
        quest including claims and description on 7 March 2023.
        This request replaced all requests on file "*subject to
        the Board of Appeal remitting the case to the Examining
        Division with an order for grant*".

VII.    Claim 1 of this request defines:

        A method for object re-identification comprising:
        providing a universal generative model of local
        descriptors;
        adapting the universal generative model to a first
        camera to obtain a first camera-dependent generative
        model;
        adapting the universal generative model to a second
        camera to obtain a second camera-dependent generative
        model or using the universal generative model as the
        second camera-dependent generative model;
        from a first image of a first object captured by the
        first camera, extracting a first image-level descriptor
        using the first camera-dependent generative model, said
        first image-level descriptor being a Fisher Vector;

from a second image of a second object captured by the
second camera, extracting a second image-level
descriptor using the second camera-dependent generative
model, said second image-level descriptor being a
Fisher Vector;
computing a similarity between the first image-level
descriptor and the second image-level descriptor; and
outputting information based on the computed
similarity, wherein:

> if the computed similarity at least meets a
> threshold, the  information comprises an indication
> that the first object and the second object match;
> or
> if the computed similarity is lower than the
> threshold, the information comprises an indication
> that the first object and the second object do not
> match;

wherein at least one of the adapting the universal
generative model to the first and second cameras,
extracting the first and second image-level descriptors
and the computing of the similarity is performed with a
computer processor;
wherein the adapting of the universal generative model
to the first and second cameras comprises extracting
local descriptors from images captured by the first and
second cameras, the local descriptors from the images
captured by the first camera being used to adapt the
universal generative model to the first camera, the
local descriptors from the images captured by the
second camera being used to adapt the universal
generative model to the second camera; wherein the
universal generative model is a Gaussian Mixture Model
and the adapting of the universal generative model to
each camera comprises a two-step Expectation-
Maximization iterative process, each iteration of said
two-step Expectation-Maximization iterative process

comprising, for each camera-dependent generative
model:
  (1) computing, for each Gaussian in the Gaussian
  Mixture Model, an estimate of the number of $n_k$
  points assigned to said Gaussian, an estimate of
  the mean $m_k$ of all points assigned to said
  Gaussian, and an estimate of the variance $s_k$ of all
  points assigned to said Gaussian; and
  (2) updating the mixture weight, mean vector, and
  covariance matrix of each Gaussian using said
  estimated number of points assigned to said
  Gaussian, the estimated mean of all points assigned
  to said Gaussian, and the estimated variance of all
  points assigned to said by:
    for the mixture weight:

$$\hat{\pi}_k^c = \left[ \tau_k^\pi \frac{n_k}{N_c} + (1 - \tau_k^\pi)\pi_k \right] \alpha$$

    for the mean vector:

$$\hat{\mu}_k^c = \tau_k^\mu m_k + (1 - \tau_k^\mu)\mu_k$$

    for the covariance matrix:

$$\hat{\sigma}_k^c = \tau_k^\sigma s_k + (1 - \tau_k^\sigma)(\sigma_k^2 + \mu_k^2) - (\mu_k^c)^2$$

  wherein $\pi_k$, $\mu_k$, and $\sigma_k$ are respectively the weight,
  mean vector, and covariance matrix of the $k$-th
  Gaussian, $\tau_k^\rho$ are the adaptation parameters for
  each parameter $\rho \in \{\pi, \mu, \sigma\}$ given by  $\tau_k^\rho = n_k/(n_k$
  $+ r^\rho)$ where $r^\rho$ is a design parameter, $N_c$ is the
  number of local descriptors extracted from images
  captured by the corresponding camera, and $\alpha$ is a
  parameter recomputed over all Gaussians such that
  $\sum_k \pi_k^c$  = 1 holds true;
  wherein $\hat{\pi}_k^c$, $\hat{\mu}_k^c$ and $\hat{\sigma}_k^c$ are respectively the
  weight, mean vector, and covariance matrix of the $k$
  -th Gaussian of the camera-dependent generative
  model.

**Reasons for the Decision**

*The application*

1.      The application relates to a method of image
        processing. The context is that of re-identification,
        i.e. the matching of an object in an image to a
        previously seen object, such as a vehicle in traffic
        surveillance (paragraphs 2 and 3).

1.1     These images may come from different cameras operating
        under different imaging conditions, for instance
        lighting. Image representations may be "shifted" as a
        function of these varying conditions (paragraphs 3-5
        and 27). To account for this, the application proposes
        to start from a universal generative model in the form
        of Gaussian Mixture Model (GMM) and to adapt it
        (through e.g. a MAP adaptation process) to individual
        cameras (paragraphs 27, 36-38, and 71-101).

1.2     As opposed to building dedicated models for each camera
        from scratch, this has the advantage that the different
        components of the generative model (the Gaussians) are
        in known correspondence. An image representation based
        on the generative model and maintaining this component
        correspondence, like the Fisher Vectors, allows the
        image matching to be performed in a simple manner,
        component by component (paragraph 82).

1.3     Results are provided in a context of license plate
        matching using images from different cameras in parking
        areas (paragraphs 106-118).

*The prior art*

2.      Document D3 describes a method wherein a GMM universal
        model is adapted to a first and a second image (using
        MAP adaptation), and then the two probability models
        are compared component by component to derive a simila-
        rity measure between the two images (paragraphs 15
        to 17, 26). This similarity metric is used for object
        classification purposes (paragraphs 40 to 44).

3.      Document D4 discusses the Fisher Vector image represen-
        tation. Inter alia, it explains that this representa-
        tion, defined on the basis of the gradient of the like-
        lihood of the test data (image patch statistics) given
        the GMM, bears a close relationship with the MAP
        adaptation procedure of a GMM to a given single image
        (section 2.5, around equations 35 to 37).

*Objections as to support and added matter*

4.      The Examining Division had raised (decision point 3.2)
        an objection of a lack of support and one of added
        matter in respect of auxiliary request 1 underlying its
        decision, stating that "*the solution to the problem of
        object matching (re-identification) announced in par.2
        requires image similarity computation*", for which the
        Fisher Vector representation was essential, and further
        that "*in technical terms, the only combination that the
        skilled person reads directly and unambiguously as
        having the essential features of the invention is that
        of MAP adaptation and FVs.*"

5.      The Board's objections as to a lack of support were
        essentially the same.

6.      The current claim 1 defines image similarity
        computation on the basis of Fisher Vectors and MAP
        adaptation of the GMM models. The objections
        therefore no longer apply.

7.      The Board found added matter in the definition of the
        MAP adaptation, considering the (previously) claimed
        generalizations to lack basis in the description. This
        objection has been overcome by amendment in the last
        (current) set of claims.

*Other objections under Article 84 EPC*

8.      In its last two communications, the Board further
        raised a number of objections under Article 84 EPC
        regarding the dependent claims (conciseness and
        clarity) and the description, which required
        adaptation. These have been overcome by amendment.

*Novelty*

9.      In establishing the equivalence with D3 (decision 1.1.
        to 1.5), the Examining Division considered that the
        adaptation to a first and second camera was anticipated
        by the application of the teachings of D3 to a video
        object. Further it considered that the first and second
        image might be ones already used during adaptation, and
        that the image descriptor was the result of the
        adaptation, i.e. the mixture model. Under this
        interpretation "*a difference in concrete technical
        terms*" could not be identified on the basis of the
        further step of extracting image-level descriptors
        using the adapted models (decision 1.4).

10.     The Board does not find this to be a technically
        reasonable interpretation of the claim.

10.1    For the skilled person, the claim recites a standard
        sequence comprising a "training" phase, when the models
        are adapted, and a "use" phase, when the already
        adapted models are used to provide descriptors for test
        images. This is clear from the wording of the claim, by
        itself and in the context of the description (see
        figure 2A). Thus, even if the adapted generative models
        themselves may be considered image (or video)
        descriptors, a further step of extracting an image-
        level descriptor from an image using the adapted models
        is not derivable from D3.

10.2    At least for this reason the subject matter of claim 1
        is new in view of D3.

*Inventive step*

*Technical effect*

11.     In its preliminary opinion, in view of the requests on
        file at that time, the Board indicated that a technical
        effect, in the sense of obtaining predictable results
        for image matching could not be acknowledged.

11.1    In particular, although the description disclosed evi-
        dence that object matching was improved in the specific
        context of license plate identification, the claims
        could not profit from this disclosure since they were
        not limited to this context. The claims only specified
        measuring image similarity, which purpose was vague
        (images may be deemed similar for vastly different
        reasons, e.g. similar luminosity, or, as here, contai-
        ning the same objects), and could not be said to serve
        a technical purpose in the absence of any claimed
        further technical use.

12.     Current claim 1 defines a method for the *re-identifica-*
        *tion* of *objects* captured by image cameras. The Board
        considers this to be a technical purpose, because it is
        tantamount to an objective measurement in physical
        reality: is the object observed now the same as the one
        observed earlier?

13.     It remains to be appreciated whether the claimed method
        provides a technical effect over substantially the
        whole scope of the claim (see G 1/19, reasons 82).

13.1    The claim defines that two objects, based on two images
        taken by two respective cameras, match if their simila-
        rity, computed using Fisher Vectors descriptors from
        adapted GMMs, is above a threshold.

13.2    The description explains that the method aims at com-
        pensating for different imaging conditions, by reflec-
        ting the shift in conditions (e.g. illuminations or
        camera angles) by a corresponding shift in the adapted
        GMM models (paragraph 27). The description also shows
        that in a case concerning license plate re-identifica-
        tion in parking facilities using in-house datasets the
        MAP adaptation claimed provides an improvement over a
        baseline no-adaptation method (paragraphs 111 to 116).
        The Board has no reason to doubt these results.

13.3    The theoretical assumptions appear sound to the Board.
        Variations in illumination and geometry should be cap-
        tured by the mixture model, and this is confirmed by
        the license plates example. In principle, the method
        should also provide improved results for other object
        matching cases, where the imaging conditions between
        two cameras differ in a similar manner.

13.4    The claimed method will not "work" under all imaginable circumstances. It is probably safe to say that no computer vision method does. For instance, the present method may fail to re-identify objects largely changing appearance. However, the skilled person will understand, from the present claims and the description, the kind of situations and its parameters (such as illumination and geometry) for which the method is designed. The method credibly works over that range of situations.

13.5    In the Board's judgment, this is sufficient to satisfy the requirement that, in the present case, a technical effect is present over substantially the whole scope of the claims (see again G 1/19, reasons 82).

*Obviousness*

14.     The Examining Division acknowledged in regard of the second auxiliary request underlying the decision, see point 5, that an additional step of extracting Fisher Vectors was not disclosed in D3, but stated that it would be obvious in the context of D3 to add a step of extracting Fisher Vectors as an image descriptor.

15.     In this respect the Board agrees with the appellant that, starting from D3, even in view of D4, it is not obvious to add a step of Fisher Vector extraction.

15.1    This is because, in D3, the adapted GMM themselves are considered to accurately describe one image. The Fisher Vectors are described in D4 rather as replacements to the image based MAP adaptation previously developed, because they capture the same information as that

obtained by adaptation starting from the same GMM (see point 3 above).

15.2    In D3, the MAP adaptation provides the object descriptor; in the application, the MAP adaptation provides camera invariance for the model (encoding the shift in conditions), and the single image object descriptor is provided by the Fisher Vectors derived from the model. Such use of adaptation to provide camera invariance is neither described, nor hinted at, by D3.

15.3    Thus, in the Board's view, the skilled person, starting from D3, may have considered Fisher Vectors to measure image similarity for classification purposes instead of the MAP adaptation method, but would not combine the two in the claimed manner.

15.4    Thus the claimed matter is not obvious in view of the prior art at hand.

16.     In view of the foregoing, the Board concludes that the subject matter of claim 1 involves an inventive step in the sense of Article 56 EPC.

**Order**

**For these reasons it is decided that:**

1. The decision under appeal is set aside.

2. The case is remitted to the first instance with the order to
   grant a patent on the basis of the following documents:

   a) claims 1-4 as filed on 7 March 2023;
   b) description pages 1-32 as filed on 7 March 2023;
   c) drawings sheets 1-6 as originally filed.


The Registrar:                              The Chairman:


L. Stridde                                  Martin Müller


Decision electronically authenticated