EUROPEAN PATENT OFFICE

U.S. PATENT AND TRADEMARK OFFICE

CPC NOTICE OF CHANGES 1499

DATE: AUGUST 1, 2023

PROJECT MP12101

The following classification changes will be effected by this Notice of Changes:

Action	Subclass	Group(s)
SCHEME:		
Notes Modified:	G06V	Subclass
DEFINITIONS:		
Definitions New:	G06F	18/00, 18/10, 18/20, 18/211, 18/2115,
		18/2133, 18/214, 18/22, 18/23, 18/24,
		18/25, 18/27, 18/40
	G06F	2123/00

No other subclasses/groups are impacted by this Notice of Changes.

This Notice of Changes includes the following [Check the ones included]:

1. CLASSIFICATION SCHEME CHANGES

- A. New, Modified or Deleted Group(s)
- B. New, Modified or Deleted Warning(s)



- C. New, Modified or Deleted Note(s)
- D. New, Modified or Deleted Guidance Heading(s)

2. DEFINITIONS

- A. New or Modified Definitions (Full definition template)
- B. Modified or Deleted Definitions (Definitions Quick Fix)
- 3. REVISION CONCORDANCE LIST (RCL)
- 4. CHANGES TO THE CPC-TO-IPC CONCORDANCE LIST (CICL)
- 5. CHANGES TO THE CROSS-REFERENCE LIST (CRL)

DATE: AUGUST 1, 2023

PROJECT MP12101

1. CLASSIFICATION SCHEME CHANGES

C. <u>New</u>, Modified or Deleted Note(s)

SUBCLASS G06V - IMAGE OR VIDEO RECOGNITION OR UNDERSTANDING

<u>Type</u> *	Location	Old Note	<u>New/Modified Note</u>
М	G06V	 This subclass <u>covers</u>: methods or arrangements for pattern recognition or machine learning specially adapted for images or video. In this subclass, the following terms or expressions are used with the meaning indicated: "pattern recognition" means detection, categorisation, authentication and identification of patterns for explanatory purposes or to derive a certain meaning in images or video, by acquiring, preprocessing or extracting distinctive features and matching, clustering or classifying these features or representations thereof; "feature extraction" means deriving descriptive or quantitative measures from images or video; "clustering" means grouping or separating patterns according to their (dis)similarity or closeness; "classification" means the identification of an object/feature as belonging to a class of objects/features by assigning of a label. In this subclass, subject matter classified in groups G06V20/00 - G06V40/00 is also classified in groups G06V10/10 or G06V10/20 respectively, if recognition relies on specific processing at the stages of acquisition or preprocessing. 	 This subclass <u>covers</u>: methods or arrangements for pattern recognition or machine learning specially adapted for images or video. In this subclass, the following terms or expressions are used with the meaning indicated:

*N = new note, M = modified note, D = deleted note

DATE: AUGUST 1, 2023

PROJECT MP12101

NOTE: The "Location" column only requires the symbol PRIOR to the location of the note. No further directions such as "before" or "after" are required.

DATE: AUGUST 1, 2023

PROJECT MP12101

2. A. DEFINITIONS (new)

G06F 18/00

Definition statement

This place covers:

Pattern recognition techniques in general.

These include in particular:

- Pre-processing, data cleansing;
- Analysing;
- Post-processing;
- Software arrangements specially adapted for pattern recognition;
- Pattern recognition of signals other than images or video.

Note:

Pattern recognition refers to the automated recognition of patterns and regularities in data and includes using statistical or machine learning techniques to solve problems such as authentication, identification, classification or clustering.

Relationships with other classification places

Systems and methods for retrieving digital information stored in databases, data repositories or file systems, locally or remotely, retrieving structured data, semi-structured data, text, audio, image, video or multimedia data from databases, or for managing document libraries are classified in group G06F 16/00.

Higher-level interpretation and recognition of images or videos, which includes pattern recognition, pattern learning and semantic interpretation as fundamental aspects, are classified in subclass G06V. These aspects involve the detection, categorisation, identification and authentication of image or video patterns by acquiring, pre-processing, extracting distinctive features or matching, supervised or unsupervised clustering or classification of these features or representations derived from them leading to one or several decisions, related confidence values (e.g. probabilities) or classification/clustering labels, for explanatory purposes or to derive a certain meaning.

Computing systems where the computation is not based on a traditional mathematical model or computer, for example neural network architectures, are classified in subclass G06N. In particular, techniques for machine learning are covered by group G06N 20/00.

DATE: AUGUST 1, 2023

PROJECT MP12101

References

Application-oriented references

Examples of places where the subject matter of this place is covered when specially adapted, used for a particular purpose, or incorporated in a larger system:

Data processing for commercial or financial purposes	G06Q
Image or video recognition or understanding	G06V
Bioinformatics	G16B
Chemoinformatics and computational material science	G16C
Healthcare informatics	G16H

Informative references

Attention is drawn to the following places, which may be of interest for search:

Control or regulating systems in general	G05B
Content-based image retrieval	G06F 16/50
Fourier, Walsh or analogous domain transformations	G06F 17/14
Matrix or vector computation	G06F 17/16
Security arrangements for protecting computer systems against	G06F 21/00
unauthorised activity	
Authentication of user input in security arrangements for	G06F 21/31
computers	
Computer-aided design	G06F 30/00
Handling natural language data	G06F 40/00
Speech recognition	G10L 15/00
Speaker recognition	G10L 17/00
Secret or secure communication	H04L 9/00

Special rules of classification

This main group covers functional aspects of pattern recognition, i.e. algorithms or computer systems specifically designed therefor, regardless of a particular application and, therefore, of any particular type of data. In this group, it is desirable to add the indexing codes of group G06F 2123/00.

Glossary of terms

In this place, the following terms or expressions are used with the meaning indicated:

DATE: AUGUST 1, 2023

PROJECT MP12101

Authentication	Verifying the identity of a sample using a test of genuineness. It generally involves a one-to-one comparison with the genuine (authentic) sample.
Categorisation	Assigning a data sample to a class according to certain distinguishing properties (or characteristics) of that class, and it generally involves a one-to-many test in which one data sample is compared with the characteristics of several classes.
Classification	Assigning labels to patterns.
Clustering	Grouping data samples in groups or classes based on their properties (or characteristics) and it generally involves a many-to- many (dis)similarity test. Grouping or separating patterns according to their closeness or dissimilarity.
Feature extraction	Deriving descriptive or quantitative measures from data.
Identification	In the context of collecting of data samples, identification means selecting a particular sample having a (predefined) characteristic which distinguishes it from the others. Several data samples are generally matched against the characteristic sample in a many-to- one process.
Pattern patterns	Data having characteristic regularity, or a representation derived from it, having some explanatory value or being able to provide an interpretation.

G06F 18/10

Definition statement

This place covers:

Removing bad data or interpolating missing values, finding hidden correlations in the data, identifying sources of data that are the most accurate and determining which data items are the most appropriate for use in analysis.

References

Informative references

Attention is drawn to the following places, which may be of interest for search:

Improving data quality; Data cleansing, e.g. de-duplication,	G06F 16/215
removing invalid entries or correcting typographical errors	
Complex mathematical operations for evaluating statistical data	G06F 17/18

DATE: AUGUST 1, 2023

PROJECT MP12101

Single-class perspective, e.g. one-against-all classification;	G06F 18/2433
Novelty detection; Outlier detection	

G06F 18/20

Definition statement

This place covers:

Extraction of features.

Matching criteria.

Clustering techniques.

Classification techniques.

Fusion techniques.

Discovering frequent patterns.

Regression.

Determining representative reference patterns.

References

Application-oriented references

Examples of places where the subject matter of this place is covered when specially adapted, used for a particular purpose, or incorporated in a larger system:

Arrangements for image or video recognition or understanding	G06V 10/00
ICT specially adapted for biostatistics; ICT specially adapted for	G16B 40/00
bioinformatics-related machine learning or data mining, e.g.	
knowledge discovery or pattern finding	
Machine learning, data mining or chemometrics for	G16C 20/70
chemoinformatics	
ICT specially adapted for computer-aided diagnosis, e.g. based	G16H 50/20
on medical expert systems	

DATE: AUGUST 1, 2023

PROJECT MP12101

Informative references

Attention is drawn to the following places, which may be of interest for search:

Image analysis G06T 7/00		Image analysis	G06T 7/00
--------------------------	--	----------------	-----------

G06F 18/211

Definition statement

This place covers:

Feature selection by using evolutionary computational techniques, e.g. genetic algorithms.

Feature selection by ranking or filtering the set of features, e.g. using a measure of variance or of feature cross-correlation.

Feature selection by evaluating class separability based on distance or (dis)similarity measures between (sub)sets of data samples. The intention for adopting distance metrics defined on features is to embed objects belonging to the same class, i.e. small intraclass distances, while embedding objects of different classes far away from each other in different classes, i.e. large interclass distances.

The processing can be iterative and, starting with no samples, it can successively include data samples by evaluating a model fit criterion, i.e. "forward" selection, or starting with all the samples and successively delete samples by evaluating a model fit criterion, i.e. backward elimination.

DATE: AUGUST 1, 2023

PROJECT MP12101

G06F 18/2115

Definition statement

This place covers:

Finding a set of features that are effective for classification, and determining whether the resulting features or feature dimensions are the most powerful for classification, e.g. such as class separability.

References

Application-oriented references

Examples of places where the subject matter of this place is covered when specially adapted, used for a particular purpose, or incorporated in a larger system:

Organisation of the matching processes	G06V 10/75
Organisation of the matching processes for character recognition or	G06V 30/19
recognising digital ink or document-oriented image-based pattern	
recognition	

G06F 18/2133

Definition statement

This place covers:

The matrix factorisation to ensure the non-negativity of the elements in the obtained matrix factors and, at the same time, nonlinear dimensionality reduction, i.e. non-negative factorisation, in order to prevent the negative elements of the selected feature subset from being contrary to the physical reality.

References

Informative references

Attention is drawn to the following places, which may be of interest for search:

DATE: AUGUST 1, 2023

PROJECT MP12101

G06F 18/214

Definition statement

This place covers:

Techniques for generating training patterns which are generally intended to improve the performance of a pattern recognition system in cases where the data representative of a certain observation/event is scarce. This can be achieved by generating new training sets by sampling the original data set, replacement of some of its data or creation of new data, as in bootstrapping.

References

Application-oriented references

Examples of places where the subject matter of this place is covered when specially adapted, used for a particular purpose, or incorporated in a larger system:

Generating sets of training patterns; Bootstrap methods, e.g.	G06V 10/774
bagging or boosting	

Informative references

Attention is drawn to the following places, which may be of interest for search:

G06F 18/22

Definition statement

This place covers:

Matching which involves comparison of data values, combinations thereof or features derived from them, in which one entity is considered as a template pattern and the other is the input pattern (template matching). The matching process might involve transforming the patterns to accommodate for distortions.

DATE: AUGUST 1, 2023

PROJECT MP12101

The proximity measures used during matching may include classical distances, such as Euclidian distances, or more involved distances, divergences or other measures between probability density functions [PDF] or other statistical representations (e.g. mean, standard deviation, moments, kurtosis or Chi-square distance), for instance:

- Kullback-Leibler divergence;
- Mutual Information;
- Bhattacharyya distance;
- Hamming distance;
- Earth mover, Wasserstein distance;
- Chi-square distance;
- Hellinger distance.

References

Application-oriented references

Examples of places where the subject matter of this place is covered when specially adapted, used for a particular purpose, or incorporated in a larger system:

Image or video pattern matching; Proximity measures in feature	G06V 10/74
spaces	
Organisation of the matching processes	G06V 10/75
Organisation of the matching processes for character recognition or	G06V 30/19
recognising digital ink or document-oriented image-based pattern	
recognition	

G06F 18/23

Definition statement

This place covers:

Techniques of grouping patterns together in order to reveal a certain structure or a meaning. These techniques aim at identifying different groups of similar entities by assigning them to a group (cluster) according to a measure of their similarity or, on the opposite, identify dissimilar items to split them into different groups. Separability is determined by measuring the similarity/dissimilarity, mostly to minimise the intra-cluster variations while maximising the inter-cluster ones. This is usually performed in a high-dimensional feature space constructed by extracting features, but can also be performed in the original domain.

DATE: AUGUST 1, 2023

PROJECT MP12101

Any pattern may belong exclusively to a single cluster (hard clustering) or it may belong simultaneously to more than one cluster up to a certain degree (fuzzy clustering) according to a proximity measure that may determine similarity or dissimilarity. Depending on the clustering method used, proximity may be defined (a) between vectors, (b) between a vector and a set of vectors (or a cluster), and (c) between sets of vectors (or different clusters).

Examples of proximity measures are: dissimilarity measures (based on I_1 , I_2 , and I_{∞} norms), similarity measures (inner product, cosine, Pearson's correlation coefficient, Tanimoto distance, etc.).

Different clustering algorithms include:

a) clustering based on statistical measures (which mainly employ numerical data) which adopt a cost function J related to possible groupings which is subject to a global or local optimisation criterion, and return a clustering that optimises J. Examples of algorithms are:

- Hard clustering algorithms, where a vector belongs exclusively to a specific cluster, e.g. k-means, k-medoids, Linde-Buzo-Gray, ISODATA, DBSCAN or Neural Gas;
- Fuzzy clustering algorithms, where a vector belongs to a specific cluster up to a certain degree, e.g. fuzzy c-means, Adaptive Fuzzy C-Shells [AFCS], Fuzzy C Quadric Shells [FCQS] or Modified Fuzzy C Quadric Shells [MFCQS];
- Probabilistic clustering algorithms, which follow Bayesian classification arguments and in which each vector is assigned to the cluster according to a probabilistic set-up, e.g. Expectation-Maximization [EM], Gaussian Mixture Model [GMM] or Mean-Shift;

b) graph-based clustering, e.g. minimum spanning tree [MST] clustering, clustering based on directed trees, spectral clustering or graph-cut optimisation;

c) Competitive learning algorithms for clustering in which a set of representatives is selected and the goal is to move each of them to regions of the vector space that are "dense" in terms of other vectors. Examples are leaky learning algorithm, Self-Organizing Maps [SOM] or Learning Vector Quantization [LVQ].

Hierarchical clustering is one of the popular techniques from the class of graph-based clustering, with its agglomerative or divisive variants. Various criteria can be used for determining the groupings, such as those based on matrix theory involving dissimilarity matrices. Algorithms included in this scheme are:

- Single link algorithm;
- Complete link algorithm;

DATE: AUGUST 1, 2023

PROJECT MP12101

- Weighted Pair Group Method Average [WPGMA];
- Unweighted Pair Group Method Average [UPGMA];
- Weighted Pair Group Method Centroid [WPGMC];
- Ward or minimum variance algorithm.

References

Application-oriented references

Examples of places where the subject matter of this place is covered when specially adapted, used for a particular purpose, or incorporated in a larger system:

Information retrieval of still images; Clustering; Classification	G06F 16/55
Information retrieval of video data; Clustering; Classification	G06F 16/75
Image or video recognition or understanding using clustering,	G06V 10/762
e.g. of similar faces in social networks	

Glossary of terms

In this place, the following terms or expressions are used with the meaning indicated:

AFC	Adaptive Fuzzy Clustering
AO	Alternative Optimisation
CCM	Compatible Cluster Merging
DBSCAN	Density-Based Spatial Clustering of Applications with Noise, a non- parametric clustering algorithm which does not require specifying the number of clusters in advance.
FCSS	Fuzzy C-Spherical Shells
FCV	Fuzzy C-Varieties
FHV	Fuzzy Hyper Volume
KNN	k-Nearest Neighbour; a classification algorithm which, for a given data sample, chooses the k most similar samples from a training set, retrieves their respective class labels, and assigns a class label to the data sample by majority decision; variant: 1NN, which is KNN for k=1.
LVQ	Learning Vector Quantisation

DATE: AUGUST 1, 2023

PROJECT MP12101

G06F 18/24

Definition statement

This place covers:

Classification in general, namely identifying to which category or which set of categories (classes) a new data or some other representation originating from it belongs, on the basis of a training set of data containing observations (or instances) whose category membership is known. Often, the individual observations are analysed into a set of quantifiable properties, known as explanatory variables or features. These properties may be categorical, ordinal, integer-valued, real-valued, etc. Other classifiers perform a class assignment by comparing current observations to previous observations by means of a similarity or distance function.

A classifier can be parametric or non-parametric depending on the type of model adopted for the observations.

Classification algorithms include those:

- based on the distance between a decision surface and training patterns, e.g. support vector machines [SVM];
- based on the distance between the pattern to be recognised and some reference, where the reference can be a prototype, a centroid of samples of the same class or the closest patterns from the same class or different classes, e.g. nearest-neighbour classification;
- based on a parametric, probabilistic model, where the model uses the Neyman-Pearson lemma, likelihood ratios, Receiver Operating Characteristics [ROC], plotting the False Acceptance Rate [FAR] versus the False Rejection Rate [FRR], Bayesian classification, etc.;
- based on a graph-like or tree-like model, e.g. decision trees, random forests, etc. Examples are the Classification and Regression Trees [CART], ID3 [Iterative Dichotomiser 3], C4.5, etc.

When the decision surface of the classifier is considered, this can be a linear classifier or non-linear classifier. Linear classifiers model the boundaries between different classes in the feature space as hyperplanes. Non-linear classifiers use e.g. quadratic, polynomial, or hyperbolic functions instead.

References

Application-oriented references

DATE: AUGUST 1, 2023

PROJECT MP12101

Examples of places where the subject matter of this place is covered when specially adapted, used for a particular purpose, or incorporated in a larger system:

Information retrieval of still images; Clustering; Classification	G06F 16/55
Information retrieval of video data; Clustering; Classification	G06F 16/75
Image or video recognition or understanding using classification	G06V 10/764

Glossary of terms

In this place, the following terms or expressions are used with the meaning indicated:

CART	Classification and Regression Tree
C4.5	Classification algorithm using a decision tree
FAR	False Acceptance Rate
FRR	False Rejection Rate
Gini impurity	A measure of how often a randomly chosen element from the set would be incorrectly labeled if it was randomly labeled according to the distribution of labels in the subset; usually used at the level of the nodes of tree-based classifiers.
ID3	Iterative Dichotomiser 3, a precursor of C4.5
ROC	Receiver Operating Characteristics

G06F 18/25

Definition statement

This place covers:

Combining the information from several sources in order to form a unified data representation.

A simple fusion process combines raw data from several sources or different modalities (e.g. fusing spatial and temporal data). Besides fusing the raw data, it is also possible to first process the sensor data to extract features and then combine the extracted features into a joint feature vector. Alternatively, it is possible to fuse classification results, e.g. inputting the features from different sensor modalities to separate classifiers, receiving respective classification scores from each classifier, and combining the individual scores into a final classification result.

Examples are probabilistic fusion, statistic fusion, fuzzy reasoning fusion, fusion based on evidence and belief theory, e.g. Dempster-Shafer or fusion by voting.

DATE: AUGUST 1, 2023

PROJECT MP12101

Fusion can also be applied at different stages of a recognition system for different purposes, e.g. for dimensionality reduction, computing robustness, improving precision and certainty in the classification decisions, etc.

References

Application-oriented references

Examples of places where the subject matter of this place is covered when specially adapted, used for a particular purpose, or incorporated in a larger system:

Image or video recognition or understanding using fusion, i.e.	G06V 10/80
combining data from various sources at the sensor level,	
preprocessing level, feature extraction level or classification	
level	

Glossary of terms

In this place, the following terms or expressions are used with the meaning indicated:

Dempster-Shafer	General framework for reasoning with uncertainty which combines	
	evidence from different sources and arrives at a degree of belief	
	(represented by a mathematical object called belief function) that	
	takes into account all the available evidence.	

G06F 18/27

Definition statement

This place covers:

Pattern recognition techniques based on regression, i.e. statistical techniques for estimating the relationships between a dependent variable (often called the "outcome" or "response" variable) and one or more independent variables (often called "predictors", "covariates" or "explanatory variables"), where the variables model the underlying data.

Common forms of regression are:

• Linear regression - the model specification is that the dependent variables are a linear combination of the parameters (but need not be linear in the independent variables). The goal is to find a line (or a more complex linear combination) that

DATE: AUGUST 1, 2023

PROJECT MP12101

most closely fits the data according to a specific mathematical criterion (e.g. by minimising the least-mean-squares criterion). For example, the method of ordinary least squares computes the unique line (or hyperplane) that minimises the sum of squared differences between the true data and that line (or hyperplane);

• Non-linear regression, e.g. polynomial, binomial, binary, logistic, multinomial logistic, etc.

References

Application-oriented references

Examples of places where the subject matter of this place is covered when specially adapted, used for a particular purpose, or incorporated in a larger system:

Image or video recognition or understanding using regression,	G06V 10/766
e.g. by projecting features on hyperplanes	

Informative references

Attention is drawn to the following places, which may be of interest for search:

Complex mathematical o	perations for evaluating statistic	al data	G06F 17/18

Glossary of terms

In this place, the following terms or expressions are used with the meaning indicated:

LMS	Least Mean Squares
RANSAC	RANdom SAmple Consensus – an iterative algorithm for fitting a linear
	mathematical model such as a line or a plane through a set of points
	by eliminating the influence of outliers

G06F 18/40

Definition statement

This place covers:

Details of software arrangements specifically adapted for pattern recognition or pattern recognition of signals. These include specifically designed user interfaces or toolboxes for solving pattern recognition problems.

DATE: AUGUST 1, 2023

PROJECT MP12101

References

Informative references

Attention is drawn to the following places, which may be of interest for search:

Interaction techniques based on graphical user interfaces [GUI]	G06F 3/048
Execution arrangements for user interfaces	G06F 9/451
GUI for database retrieval, e.g. presentation of query results,	G06F 16/248, G06F
browsing or visualisation therefor	16/26
ICT specially adapted for bioinformatics-related data	G16B 45/00
visualisation, e.g. displaying of maps or networks	
Data visualisation for chemoinformatics	G16C 20/80

G06F 2123/00

Definition statement

This place covers:

Data types on which pattern recognition is applied such as time-series data.

References

Application-oriented references

Examples of places where the subject matter of this place is covered when specially adapted, used for a particular purpose, or incorporated in a larger system:

Analysing medical signals, e.g. bioelectric signals or blood	A61B 5/00
pressure	
Processing radar or similar signals	G01S
Analysis of chromatographic signals	G01N 30/86
Processing seismic signals	G01V 1/28
Transmission systems for measured values, control or similar	G08C
signals	