

Discovering the Discriminating Power in Patient Test Features Using Visual Analytics: A Case Study in Parkinson's Disease

Panagiotis Moschonas¹(✉), Elias Kalamaras¹, Stavros Papadopoulos¹,
Anastasios Drosou¹, Konstantinos Votis¹, Sevasti Bostantjopoulou²,
Zoe Katsarou², Charalambos Papaxanthis³,
Vassilia Hatzitaki⁴, and Dimitrios Tzovaras¹

¹ Information Technologies Institute, Centre for Research and Technology Hellas,
Thessaloniki, Greece

{moschona,kalamar,spap,drosou,kvotis,Dimitrios.Tzovaras}@iti.gr

² Department of Neurology, Aristotle University of Thessaloniki,
Thessaloniki, Greece

³ Universit de Bourgogne, UFR STAPS, Campus Universitaire, Dijon, France

⁴ Department of Physical Education and Sports Sciences,
Aristotle University of Thessaloniki, Thessaloniki, Greece

Abstract. This paper presents a novel methodology for selecting the most representative features for identifying the presence of the Parkinson's Disease (PD). The proposed methodology is based on interactive visual analytic based on multi-objective optimisation. The implemented tool processes and visualises the information extracted via performing a typical line-tracking test using a tablet device. Such output information includes several modalities, such as position, velocity, dynamics, etc. Preliminary results depict that the implemented visual analytics technique has a very high potential in discriminating the PD patients from healthy individuals and thus, it can be used for the identification of the best feature type which is representative of the disease presence.

Keywords: Parkinson's disease · Visual analytics · Multi-objective optimisation · Feature discrimination power

1 Introduction

Parkinson's disease (PD) is a degenerative disorder of the central nervous system that is mainly affecting the motor system. Its most obvious symptoms are movement-related, including shaking, rigidity, slowness of movement and difficulty with walking and gait [8]. Tracking in a correct way the progress of the

disease is crucial for the quality of patient's life, thus the need of an objective and consistent way of measuring the patient motor dexterity is essential.

There are several established scales for measuring the progress of the disease. The Unified Parkinson's Disease Rating Scale [16] is the most commonly used metric for clinical study and is used as a severity rating method. An older scaling method known as the Hoehn and Yahr scale [8], and a similar scale known as the Modified Hoehn and Yahr scale, have also been commonly used. The later defines five basic stages of progression. These scale measurement are derived from qualitative questionnaires asked by doctors, thus they enclose the danger of subjectivity in their results: different clinical examinations may produce different answers. Additionally, the repetition period of such solutions are in the class of several months.

Electronic devices from various domains have been also used for measuring the PD progress. The current market offers a variety of wristbands capable of measuring the tremor [6]. However, these are used for complimentary assessment to the typical clinical examination. Moreover, their effectiveness is high when the tremor is present to the patient arm, but their efficiency is not guaranteed in early stages of the disease. An increasing set of smartphone and tablet PD applications has arisen in the market. Despite most of them are used for scheduling, there is a small subset which makes use of tests for tracking the progress of the disease. In most of these tests, the user has to use his/her hand to track a line or a shape and the software extract metrics representative of the current state. Factors such as velocity, target deviation, reaction time, minimum jerk, are recorded and compared to the ones measured from previous states. The results are displayed in curves, allowing the user to track the PD progress.

However, even providing the tests, there is still the matter of how the program is going to correlate the captured information with the disease presence on the individual. Comparing the velocity or deviation profiles per day measurements is not enough as it could not be representative of the specific patient symptoms [18]. For such reasons, proper visualisation capable of identifying which of the extracted features encapsulate and discriminate better the examined person status is needed. In this paper we apply a novel multi-objective visualisation for identifying which of the measured quantities are best for discriminating the presence of the Parkinson's Disease.

1.1 Relevant Work

Information visualization concerns the use of interactive computer graphics to get insight into large amounts of abstract data, such as multivariate, hierarchical and network data [23]. Conventional visualization techniques, such as bar-charts, pie-charts, and line-charts, are useful for the depiction of information of a higher level, but fail to depict large and complex data sets in detail. Therefore, and in order to enable the easy extraction of patterns, trends and outliers, a variety of novel visualization methods are continuously being developed for specific applications, e.g. [1, 4, 22].

Their ability to provide insight into large amounts of complex data make visual analytics techniques especially useful for healthcare and biomedical applications. This has resulted in numerous existing methods and tools which utilize various visualization types and user interaction levels [10, 21]. Recent works have employed visual properties such as color and position [15, 24] or animation [17], in order to visually encode patient information, group patients with similar characteristics together and discriminate between different events. Visual queries have also been utilized in combination with pattern mining and interactive visualization, in order to explore large datasets more efficiently.

The continuous development in the field of information visualization has led to a variety of new methods and techniques that enable people to understand the phenomena behind large amounts of data, and which increasingly find their way to Health and real-life applications. However, and due to their specialized nature, many of these methods have limited use outside their initially targeted application. For the majority of methods, data are considered to be homogenous and in most cases just one or two types of data are supported. In addition, data are assumed to come from a single source and to be clean and exact, whereas noisy, polluted, uncertain, and missing data are rarely dealt with. Scalability is another common issue, since it is typically limited to thousands of elements. Therefore, if one of the above assumptions is violated, standard methods from information visualization fall short, and a need for new representations arises.

Since several attributes are usually available for each patient, such as age, speed of task completion, acceleration, etc., they can be combined to provide better insights in the data. The combination of multiple sources of information for classification, clustering, visualization, etc. has generally been handled by multimodal fusion methods [2]. A first category of fusion methods simultaneously combine characteristics of all modalities, e.g. through weighted sums. In [20], a graph is constructed for each of multiple attributes and the graph Laplacians are then fused. Graph-based techniques are also used in [13], which employs Multiple Kernel Learning [9] for fusion. A second category are methods that utilize information of one modality to assist learning in another modality in an iterative manner, such as [3, 14].

A different principle is followed by works such as [11, 12], where visualization of each modality is formulated as an optimisation problem and then multi-objective optimisation techniques [5, 7] are utilized to simultaneously optimize all objectives and resulting in a set of Pareto-optimal solutions, instead of one. Although such methods demonstrate the effectiveness of multimodal fusion, for a wide range of applications, they have not been adopted for visualization of healthcare data, where the combination of multiple modalities may reveal important patient behavioural characteristics that can assist decision making. In this respect, this paper presents an application of the techniques presented in [11, 12], in the task of clustering patients according to multiple attributes.

2 Applied Methodology

A set of features is first extracted by performing a simple test on a tablet device. During the test a vertical line appears at random position, while the subject has to track it as fast as possible with a stylus pen. After the test, a set of features is extracted and passed into the visual analytics engine.

The extracted features are the following: (a) deviation in x axis from target, i.e. position in pixels, (b) velocity, (c) acceleration, and (d) simulated muscle activation, extracted using methodology described in [19]. These features are normalised in the time component. This is achieved by splitting the whole recording into events which start when the target line changes its position. Then a representative signal of each feature type is extracted by averaging the event signals. The resulted signals are fed to the Multi-objective visual analytics engine in order to analyse the discriminating power of each of the feature types.

2.1 Multi-objective Visualisation

The feature extraction procedure results in feature vectors describing various characteristics of a patient, such as age, speed of task completion, acceleration. Each of these types of features determines a specific type of similarities and dissimilarities among the users. An adequate visualisation of the patient data should be able to visualize these similarities. A common and straightforward visualisation scheme is to consider each patient as a point on the screen and using the relative position of the patients to denote similarity, by placing similar patients close to each other and dissimilar ones away from each other. However, the presence of multiple notions of similarity, due to the multiple types of features extracted from a single patient, renders the problem of visualizing similarities non-trivial, since many types of similarity must be visualized simultaneously, or some kind of compromise among the various feature types should be considered. In this respect, the multi-objective visualisation method of [11] is used hereby, which exactly considers multiple notions of similarity. The method is briefly presented in this section. The method proceeds by first considering each feature type separately, formulating visualisation as a single-objective optimisation problem, and then combining multiple objectives in a multi-objective optimisation setting.

Visualisation of a Single Feature Type: Let $\mathcal{O} = \{O_1, O_2, \dots, O_N\}$ denote a set of N patients. Each patient O_i is itself a set of M feature vectors:

$$O_i = \{\mathbf{o}_{i,1}, \mathbf{o}_{i,2}, \dots, \mathbf{o}_{i,M}\}, \mathbf{o}_{i,m} \in \mathbb{R}^{D_m}.$$

Each feature vector $\mathbf{o}_{i,m}$ is a vector of dimension D_m , which is specific for the m -th feature type.

Considering just a single feature type m , a distance function

$$d_m : \mathbb{R}^{D_m} \times \mathbb{R}^{D_m} \rightarrow \mathbb{R}$$

is used to calculate the distance between the feature vectors of two patients. Hereby, the L1 distance measure is used:

$$d_m(\mathbf{o}_{i,m}, \mathbf{o}_{j,m}) = \sum_{k=1}^{D_m} |o_{i,m,k} - o_{j,m,k}|,$$

where $o_{i,m,k}$ is the k -th component of vector $\mathbf{o}_{i,m}$. The distance measure is used to construct a complete graph $G_m(\mathcal{O}, E_m)$, $E_m \subseteq \mathcal{O} \times \mathcal{O}$, of the patients, where each patient $O_i \in \mathcal{O}$ is represented by a vertex and there is an edge $\{O_i, O_j\} \in E_m$ between every pair of patients, weighted by the corresponding distance between them. As a further step, the minimum spanning tree (MST), $T_m(\mathcal{O}, E'_m)$, $E'_m \subset E_m$, of the graph is computed, in order to reduce the number of edges, while keeping the similarity information.

The MST is then used to guide the positioning of the patient vertices on the 2-dimensional screen. Let $\mathbf{p}_i = (x_i, y_i)^T$, $x_i, y_i \in \mathbb{R}$, be the coordinates of patient i on the screen. Let also $\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N)^T$ be the matrix collecting all the 2D points as its rows. Then, the placement \mathbf{P} of the N points is calculated by minimizing the following objective function:

$$J_m(\mathbf{P}) = \sum_{i=1}^N \sum_{j=1, j \neq i}^N \frac{q^2}{\|\mathbf{p}_i - \mathbf{p}_j\|} + \sum_{i,j: (O_i, O_j) \in E'_m} k \|\mathbf{p}_i - \mathbf{p}_j\|^2.$$

Intuitively, this function evaluates the potential energy of the tree, if each vertex is considered as a charged particle and the edges as springs attached to pairs of them. The first term of the sum is the energy of the charges repelling one another, following Coulomb’s law. The second term is the energy of the attracting springs, according to Hooke’s law. If such a system of charges and springs is let to act freely, the particles will tend to repel one another, while the springs will keep together particles connected by edges, thus unfolding the tree structure. The lowest energy position is one where the tree has been completely unfolded

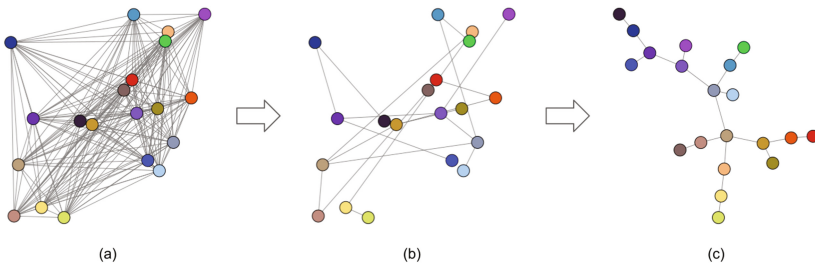


Fig. 1. Example of the potential objective minimization. (a) The points and their similarities form a complete graph. The graph’s vertices are placed at random positions. (b) The minimum spanning tree of the graph is calculated. (c) By minimizing the potential objective, the vertices are moved so that the tree’s structure is apparent. (Color figure online)

and its structure is easily visible. In this position, the patients with similar characteristics will be put close to each other, while dissimilar patients will be put away from each other. An example of the overall procedure is illustrated in Fig. 1, where the color of each point is used to visually denote the example characteristics of each point. After the minimization of the potential objective, in Fig. 1(c), the structure of the tree is apparent and similar colors are placed close to each other.

Handling Multiple Feature Types: The above procedure holds for a specific feature type m . If a different feature type is used, different distances will be calculated among the points, resulting in a different tree, a different objective function and a different ultimate positioning. The multiple feature types result in a set of objective functions $\mathcal{J}(\mathbf{P})$, instead of a single one:

$$\mathcal{J}(\mathbf{P}) = \{J_1(\mathbf{P}), J_2(\mathbf{P}), \dots, J_M(\mathbf{P})\},$$

so that the goal is to minimize them all simultaneously:

$$\mathbf{P}_{\text{opt}} = \arg \min_{\mathbf{P}} \mathcal{J}(\mathbf{P}).$$

The multiple objective J_m are conflicting, since the optimal solution for one of them is not optimal for another, thus a single solution cannot in general be reached. Such problems of conflicting objectives are handled by multi-objective optimisation techniques, which result in a set of optimal trade-offs among the objectives, namely the *Pareto set*. Multi-objective optimisation is based on the notion of *Pareto dominance* among the feasible solutions. A solution dominates another one if it has a smaller value for at least one objective and there is no objective for which it has a larger value. Between two solution, the dominant one is preferred, since it is impartially better than the other one, with respect to all objective functions, without sacrificing any of them. Formally, a solution \mathbf{P}_1 dominates another solution \mathbf{P}_2 , if

$$J_m(\mathbf{P}_1) \leq J_m(\mathbf{P}_2), \forall m \in \{1, \dots, M\}, \text{ and} \\ \exists k \in \{1, \dots, M\} : J_k(\mathbf{P}_1) < J_k(\mathbf{P}_2).$$

If two solutions mutually do not dominate each other, they are said to be *incomparable*, since there can be no impartial judgment as to which is better than the other. The goal of multi-objective optimisation is to compute the set of solutions that dominate all other feasible solutions but are mutually incomparable. This set is called the Pareto set and consists a set of optimal trade-offs among the multiple objectives. In Fig. 2, the feasible and optimal solutions for an example problem of two objectives are depicted.

The gray-shaded area represents the set of all feasible solutions, while the bold border in the lower left of the feasible area represents the solutions of the Pareto set. In such diagrams in the space of the objective function values, the

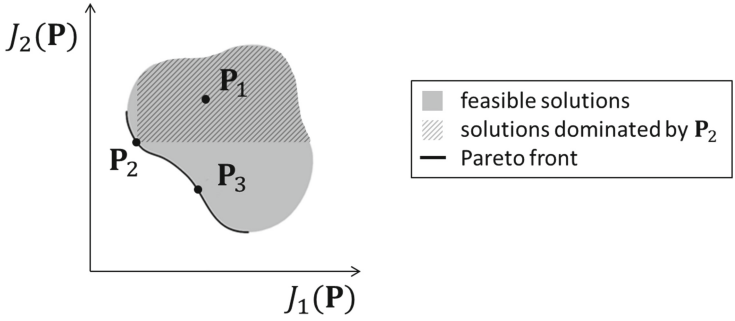


Fig. 2. Example Pareto diagram illustrating the Pareto front for a problem of two objectives, J_1 and J_2 . The gray area represents the set of all feasible solutions, while the bold border is the Pareto front. Solution \mathbf{P}_2 dominates \mathbf{P}_1 , as well as all solutions within the hatched area. Solutions \mathbf{P}_2 and \mathbf{P}_3 are incomparable.

solutions corresponding to the Pareto set are called the *Pareto front*. Three example solutions are shown. Solution \mathbf{P}_2 dominates \mathbf{P}_1 since both objectives have smaller values at \mathbf{P}_2 . Similarly, \mathbf{P}_2 dominates all solutions within the hatched area. On the other hand, solutions \mathbf{P}_2 and \mathbf{P}_3 are incomparable, since none dominates the other. All the solutions of the Pareto front are mutually incomparable, since decreasing one objective leads to increasing the other.

Computing the Pareto front, which is the goal of multi-objective optimisation, means presenting the decision maker with a minimal set of optimal trade-offs, from which to select. In this paper, this set of solutions corresponds to different trade offs among the various medical and simulated features describing the patients. By selecting among the trade-offs, through an application interface, the doctor or any decision maker can put more focus on various characteristics and thus view different kinds of relationships and groupings among the data.

3 Preliminary Results

A total of 42 PD patients of various ages (mean: 63, stdev: 8.1) and 10 healthy/control subjects of matching ages (mean: 64.9, stdev: 7.75) were measured. All PD subjects were tested with an on-medication treatment state. The subjects performed the test with both hands and the extracted features were averaged, thus resulting into one feature-set per subject. The proposed algorithm computes the visualisation output of the 52 instances almost instantly ($t < 0.05$ s). The interactive environment allows its user to re-adjust the objective weights and updates the dots position without any lag. In terms of scalability the algorithm performs at $O(N \log N)$, where N is the number of nodes. Experiments in the application domain of mobile network security have shown that a dataset of 4800 points needs less than 3s for the output solution [12]. The aforementioned experiments were performed using a 8-threaded Intel Core i7 processor running at 4 GHz.

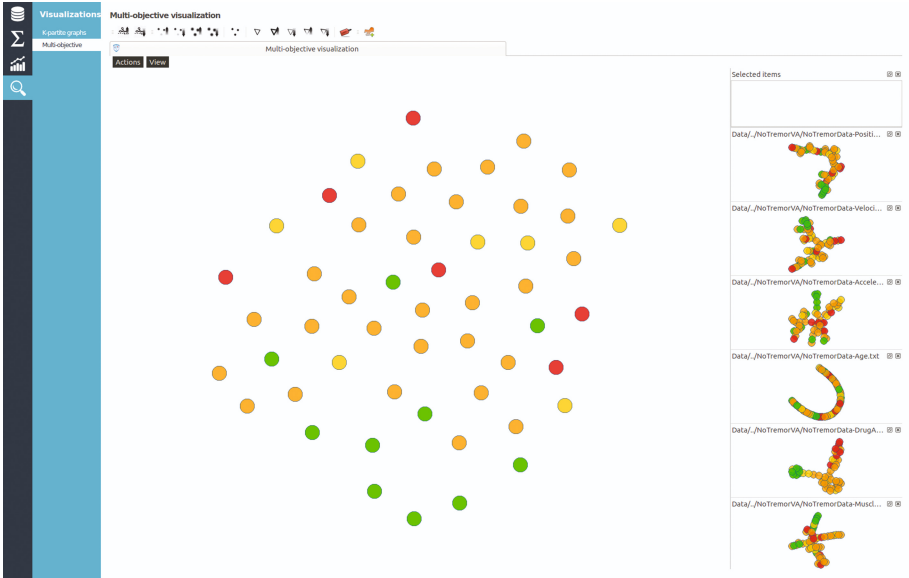


Fig. 3. Multi-objective result of all the feature type combined. The gathering of the green dots representing the control subjects in the lower part is not concentrated. (Color figure online)

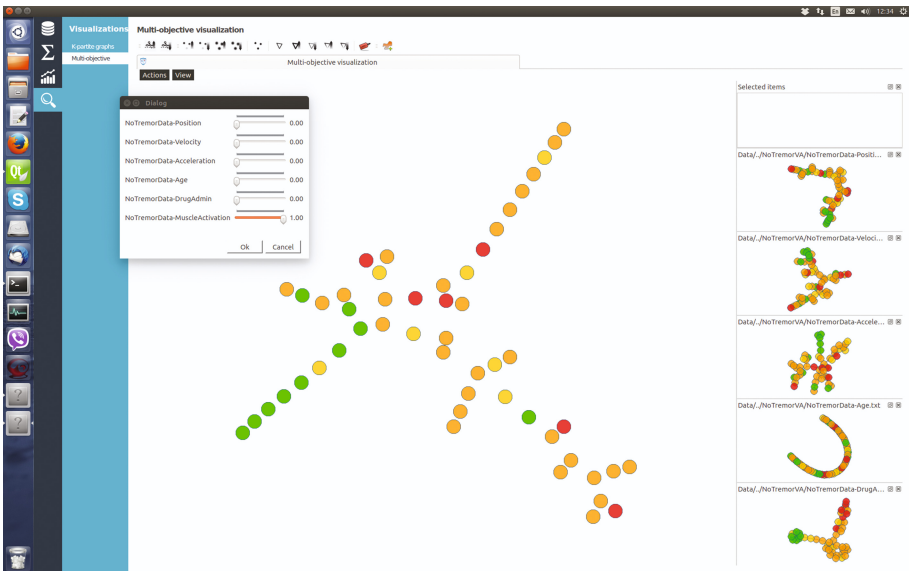


Fig. 4. Testing the impact of each feature: increasing the weight of the “Muscle Activation” feature results into a robust patient-healthy separation. (Color figure online)

The patient space of the combining all the features, using equal weighting among them, is depicted in Fig. 3. Green coloured dots represent a control subject, while yellow (mild condition), orange (medium condition) or red (severe condition) dots represent the PD patients. As it is shown, a combination of all the features results into a sparse cluster of the green dots. However, it fails to cluster three of the control subjects. This fact can be translated as that the selected combination of the feature types does not provide enough discriminating power.

By adjusting the weights among the feature types, the most crucial one for separating the healthy from the patient cohort may be found. In Fig. 4, the feature representing the extracted Muscle Activations separates the two groups in much more robust way, thus, defining it as the most significant for this type of application. Other features have been tested, but none has provided such good discrepancy.

4 Conclusion

A visual analytics methodology based on multi-objective optimisation was applied to a typical PD line tracking test output and managed to identify successfully the most crucial feature able to separate the healthy from the patient groups. In the future work, there are plans of investigating the potential of the methodology with other neurodegenerative genetic disorders, such as the Huntington's Disease.

Acknowledgement. Partial funding for this work was provided by the EU co-funded project NoTremor (EC FP7 Grant agreement no 610391).

References

1. Aigner, W., Miksch, S., Schumann, H., Tominski, C.: Visualization of Time-Oriented Data. Springer Science & Business Media, London (2011)
2. Atrey, P.K., Hossain, M.A., El Saddik, A., Kankanhalli, M.S.: Multimodal fusion for multimedia analysis: a survey. *Multimedia Syst.* **16**(6), 345–379 (2010)
3. Blum, A., Mitchell, T.: Combining labeled and unlabeled data with co-training. In: *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pp. 92–100. ACM (1998)
4. Carroll, L.N., Au, A.P., Detwiler, L.T., Fu, T.C., Painter, I.S., Abernethy, N.F.: Visualization and analytics tools for infectious disease epidemiology: a systematic review. *J. Biomed. Inf.* **51**, 287–298 (2014)
5. Coello, C., Lamont, G., Van Veldhuizen, D.: *Evolutionary Algorithms for Solving Multi-objective Problems*. Springer, New York (2007)
6. Coy, J.A., Mehrkens, J.H., Roppenecker, D.B., Lueth, T.C.: Finding the center of parkinson's disease. a novel measurement device for quantifying motor symptoms during DBS-surgery. In: *2014 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 1691–1696. IEEE (2014)
7. Ehrgott, M.: *Multicriteria Optimization*, vol. 2. Springer, Heidelberg (2005)

8. Goetz, C.G., Poewe, W., Rascol, O., Sampaio, C., Stebbins, G.T., Counsell, C., Giladi, N., Holloway, R.G., Moore, C.G., Wenning, G.K., et al.: Movement disorder society task force report on the Hoehn and Yahr staging scale: status and recommendations the movement disorder society task force on rating scales for parkinson's disease. *Mov. Disord.* **19**(9), 1020–1028 (2004)
9. Gönen, M., Alpaydm, E.: Multiple kernel learning algorithms. *J. Mach. Learn. Res.* **12**, 2211–2268 (2011)
10. Holzinger, A., Jurisica, I.: Knowledge discovery and data mining in biomedical informatics: the future is in integrative, interactive machine learning solutions. In: Holzinger, A., Jurisica, I. (eds.) *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics*. LNCS, vol. 8401, pp. 1–18. Springer, Heidelberg (2014)
11. Kalamaras, I., Drosou, A., Tzovaras, D.: A multi-objective clustering approach for the detection of abnormal behaviors in mobile networks. In: *2015 IEEE International Conference on Communication Workshop (ICCW)*, pp. 1491–1496. IEEE (2015)
12. Kalamaras, I., Papadopoulos, S., Drosou, A., Tzovaras, D.: MoVA: a visual analytics tool providing insight in the big mobile network data. In: Chbeir, R., Manolopoulos, Y., Maglogiannis, I., Alhajj, R. (eds.) *AIAI 2015. IFIP AICT*, vol. 458, pp. 383–396. Springer, Heidelberg (2015). doi:[10.1007/978-3-319-23868-5_27](https://doi.org/10.1007/978-3-319-23868-5_27)
13. Lin, Y.Y., Liu, T.L., Fuh, C.S.: Multiple kernel learning for dimensionality reduction. *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(6), 1147–1160 (2011)
14. Nigam, K., Ghani, R.: Analyzing the effectiveness and applicability of co-training. In: *Proceedings of the Ninth International Conference on Information and Knowledge Management*, pp. 86–93. ACM (2000)
15. Ordóñez, P., DesJardins, M., Lombardi, M., Lehmann, C.U., Fackler, J.: An animated multivariate visualization for physiological and clinical data in the ICU. In: *Proceedings of the 1st ACM International Health Informatics Symposium*, pp. 771–779. ACM (2010)
16. Ramaker, C., Marinus, J., Stiggelbout, A.M., van Hilten, B.J.: Systematic evaluation of rating scales for impairment and disability in parkinson's disease. *Mov. Disord.* **17**(5), 867–876 (2002)
17. Rind, A., Aigner, W., Miksch, S., Wiltner, S., Pohl, M., Drexler, F., Neubauer, B., Suchy, N.: Visually exploring multivariate trends in patient cohorts using animated scatter plots. In: Robertson, M.M. (ed.) *EHAWC 2011 and HCII 2011*. LNCS, vol. 6779, pp. 139–148. Springer, Heidelberg (2011)
18. Snow, B.: Objective measures for the progression of parkinsons disease. *J. Neurolo. Neurosurg. Psychiatry* **74**(3), 287–288 (2003)
19. Stanev, D., Moschonas, P., Votis, K., Tzovaras, D., Moustakas, K.: Simulation and visual analysis of neuromusculoskeletal models and data. In: Chbeir, R., Manolopoulos, Y., Maglogiannis, I., Alhajj, R. (eds.) *AIAI 2015. IFIP AICT*, vol. 458, pp. 411–420. Springer, Heidelberg (2015). doi:[10.1007/978-3-319-23868-5_29](https://doi.org/10.1007/978-3-319-23868-5_29)
20. Tong, H., He, J., Li, M., Zhang, C., Ma, W.Y.: Graph based multi-modality learning. In: *Proceedings of the 13th Annual ACM International Conference on Multimedia*, pp. 862–871. ACM (2005)
21. Turkay, C., Jeanquartier, F., Holzinger, A., Hauser, H.: On computationally-enhanced visual analysis of heterogeneous data and its application in biomedical informatics. In: Holzinger, A., Jurisica, I. (eds.) *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics*. LNCS, vol. 8401, pp. 117–140. Springer, Heidelberg (2014)

22. Von Landesberger, T., Kuijper, A., Schreck, T., Kohlhammer, J., van Wijk, J.J., Fekete, J.D., Fellner, D.W.: Visual analysis of large graphs: state-of-the-art and future research challenges. In: *Computer Graphics Forum*, vol. 30, pp. 1719–1749. Wiley (2011)
23. Ward, M.O., Grinstein, G., Keim, D.: *Interactive Data Visualization: Foundations, Techniques, and Applications*. CRC Press, Natick (2010)
24. Wongsuphasawat, K., Guerra Gómez, J.A., Plaisant, C., Wang, T.D., Taieb-Maimon, M., Shneiderman, B.: LifeFlow: visualizing an overview of event sequences. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1747–1756. ACM (2011)