

Collaboration in Computer Vision using Scientific Workflows

Kabir Chug, University of Texas, Austin
 Ricky J. Sethi, Fitchburg State University, rickys@sethi.org

Abstract—Collaboration, extension, and reproduction of research is of great importance in computer vision. Scientific workflows offer a unique framework for distributed collaboration and sharing of experiments. They provide a structured, end-to-end analysis methodology that easily and automatically allows for standardized replication and testing of models, inter-operability of heterogeneous codebases, and incorporation of novel algorithms. In this paper, we introduce the use of scientific workflows in computer vision to aid collaboration.

I. INTRODUCTION

Computational research projects usually involve heterogeneous implementations of algorithms written in different programming languages by many researchers who are often also dispersed geographically. Such projects are usually configured for specific runs with long-forgotten parameters and configurations, which can be a nightmare to integrate or reuse with other projects. Trying to manage and consolidate such disparate codebases written by colleagues or students, some of whom may have moved on, can hobble research thrusts. Even worse, the dissemination of research itself can be crippled by such practices since this is likely to make reproducing past results difficult.

This can lead to unreliable comparisons to others' work, insufficient verification of the state-of-the-art, inadvertent errors in publications, and a general inability to efficiently and easily extend previous work by co-authors, postdocs, and graduate students. **Scientific Workflows**, which capture an end-to-end analysis composed of individual analytic steps as a dependency graph, can aid in such collaborations. They represent complex applications as a dependency network of individual computations linked through control or data flow.

There are several workflow systems used in scientific data analysis [1] but none have been applied until now to computer vision. We use the WINGS workflow system [2], which has three key features that make workflows accessible to computer vision researchers: a simple dataflow structure, a simple web interface, and the ability to publish workflows as web objects [3]. In addition, the WINGS standard repository includes workflow fragments for general machine learning packages like Weka, CLUTO for document clustering, etc.

Users can export these workflows and make them available as part of a workflow library like the WINGS standard repository so that other researchers can directly utilize any single component (or the entire component collection) in their own workflows by simply importing those exported web objects. The exported workflows can also be adapted by adding or

changing any component. These *workflows web objects* thus allow for complete sharing of computational experiments that are represented as scientific workflows and provide for the sharing, reproducibility, and extension of the original experiments by examining or executing them online or downloading and importing them locally.

The WINGS workflow system was developed to assist scientists in managing complex computations [4], [5], [2], [6], [7] and it has been used in several large-scale distributed scientific applications, including text analytics [8], [9] and biomedical image processing [10]. WINGS uses semantic workflow representations that capture the requirements and semantic constraints of individual steps and datasets explicitly, as well as workflow reasoning algorithms to generate possible combinations of workflow components systematically and to validate them. WINGS provides a common, structured framework for composing algorithms and methods for machine learning and computer vision analysis; it is easily extensible to incorporate new algorithms and methods and to process additional types of data.

II. WORKFLOWS FOR COMPUTER VISION

In this paper, we designed Computer Vision workflows for common tasks, such as computing optical flow, Kalman Tracker, n-cuts algorithm, etc., as shown in Table I. In addition, we developed implementations of novel algorithms like the Phase Space model [11], Group Transition Ratio [12], etc. We used these in conjunction with standard Data Analytic workflows [9].

Workflows for tasks such as classification, clustering, etc., are often composed of smaller workflow fragments that preprocess data, prepare the testing/training sets, and setup the learning task [9]. We have developed several workflow fragments for computer vision tasks which can easily be incorporated in new analyses and research thrusts. These workflow fragments, or those created by users themselves, can be combined to form end-to-end workflows. Users can then export [3] these workflows and make them available as part of a workflow library like those in the WINGS standard repository; other researchers can directly utilize any single component (or the entire component collection) in their own workflows by simply importing those components. In addition, these exported workflows can be adapted for specific research by adding or changing any component.

Workflows can be exported by publishing them as web objects using Linked Data principles [3]. These web objects,

Table I: Workflow fragments created and published for computer vision applications.

Category	Workflow Fragments
Computer Vision Fragments	OpenCV components (Optical Flow, Kalman tracker, Mixture of Gaussians, Particle filter, etc.), N-Cuts, PhaseSpace, G_{tr} , RelativeVelocity, RelativeDistance, Image Extractor
General Machine Learning Fragments	K-Means, LDA, Mallet, libSVM
Statistical Evaluation Fragments	Confusion Matrices, Heatmaps, Precision-Recall Curves, ROC Curves, AUC Curves, Equal Error Rate, F-Measure

represented in RDF, allow direct access via unique URIs to workflows, their components, and their associated datasets. This allows any scientist to inspect a workflow without having to ask the original investigators for details and, in fact, to reproduce the exact results originally obtained as all parameters and details of the experiments are encapsulated in the publicly accessible web objects. These web objects can then be imported into any workflow system that is compatible with the standard Open Provenance Model for workflow publication [3].

III. COLLABORATION VIA WORKFLOWS

Scientific workflows allow collaboration between colleagues, postdocs, graduate students, undergraduate students, and even high school students. We have previously shown collaboration between high school students using standard Data Analysis workflows [9]. We continued these kinds of workflows-based collaborations and found that undergraduate students were able to collaborate on Computer Vision tasks by having one student mark the dataset with ground truth in a geographically dispersed location (Los Angeles, CA) and add it to the workflows. Another student (in Austin, TX) developed some of the workflow fragments for image analysis, including computation of entropy in an image.

Researchers in Fitchburg, MA were then able to directly utilize these components and workflows to immediately conduct an analysis for both an image analysis task and a video recognition task. The students in Los Angeles, CA were also able to add additional implementations of the Latent Dirichlet Allocation (LDA), including a parallel version as shown in Figure 1. Initial results were computed by the groups in Los Angeles, CA, Fitchburg, MA, and Austin, TX to compare to the speed-up offered by the parallelized LDA when compared to the standard LDA, as shown in Figure 2.

For the video activity recognition task, we utilized the *Group Transition Ratio*, G_{tr} , from [13] to quantitatively define a group as well as the Atomic Group Actions. The G_{tr} is defined as $G_{tr} = \frac{L}{\lambda}$, where λ is the mean free path and L is the characteristic length. The characteristic length is usually a convenient reference length that is a constant of a given configuration.

We used the G_{tr} to identify when a collection of objects can be considered as separate individuals, a group, or a crowd. In addition, we used the time variance of the G_{tr} to determine when a collection of objects transitions between being individuals, groups, or crowds. In this paper, we implemented the G_{tr} in matlab as a component in the WINGS workflows system and tested it against the Atomic Group Actions dataset [12]. The abstract workflow for the entire evaluation is shown

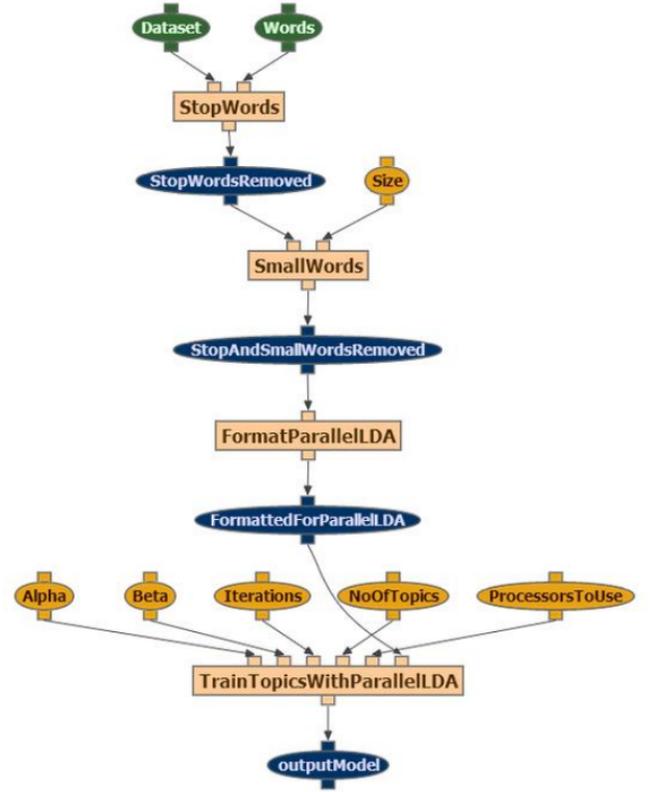


Figure 1: Workflow for the Parallelized Latent Dirichlet Allocation (LDA)

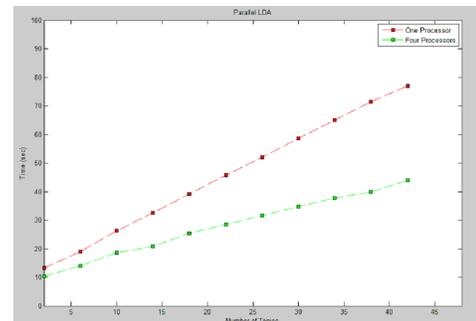
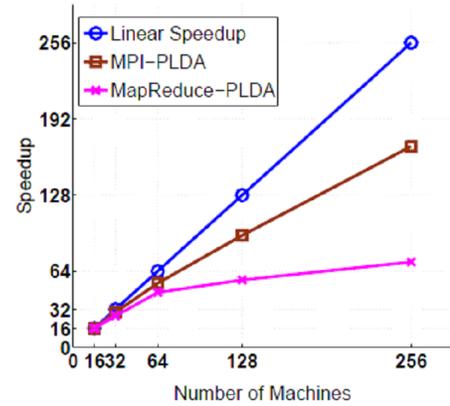


Figure 2: Comparison of the parallelized and standard LDA

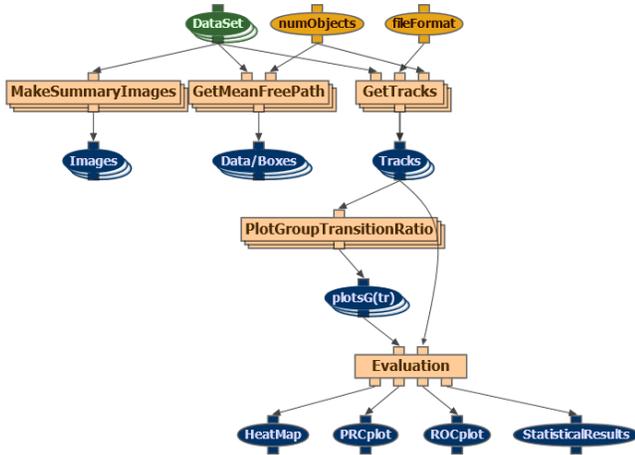


Figure 3: Abstract Workflow for the G_{tr} model using the Atomic Group Actions Dataset

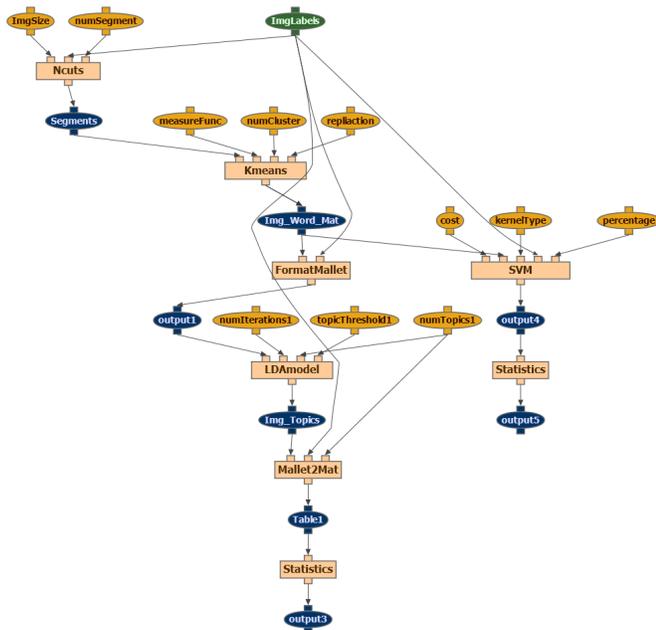


Figure 4: Abstract Workflow for the Multimedia Social Media Task

in Figure 3. The Evaluation component represented there is a composite representation of multiple statistical evaluation components from Table I. The re-use of pre-created workflow fragments like the statistical fragments shown there allowed for a rapid development cycle in addition to providing the ability to export the entire analysis and dataset as web objects for reproducibility.

We first show some qualitative results for the G_{tr} in Figure 3 that were generated automatically from the workflow. In Figure (5), we show how the G_{tr} varies with time to characterize the action in a video; results for all categories of the G_{tr} graph along with representative frames from the

video are overlaid to show the action. The thresholds were empirically determined automatically by the workflow to be $G_{tr} \lesssim 0.1$ for crowds, $0.1 \lesssim G_{tr} \lesssim 1.5$ for groups, and $G_{tr} \gtrsim 1.5$ for individuals. In Figure (5), we show how the G_{tr} decreases with time for *group formation* in (a), increases with time for *group dispersal*, and stays the same for *group movement* in (b).

We further demonstrate the collaborative benefits of scientific workflows by extending a pre-existing social media analysis task. The Human Trafficking Detection (HTD) project analyzes posts on various sites on the internet, downloading approximately 20GB of data every day, in order to determine if the subject of that post is a victim of human trafficking. The HTD developed by an original group of 1 researcher and 1 postdoc in Los Angeles, CA only looked at the text data in posts. This was extended by another group in Los Angeles, CA and Fichburg, MA composed of 2 researchers, 1 postdoc, 1 graduate student, and 2 undergraduate students to examine both the text and images associated with those posts, as well as the fusion of these two distinct data domains.

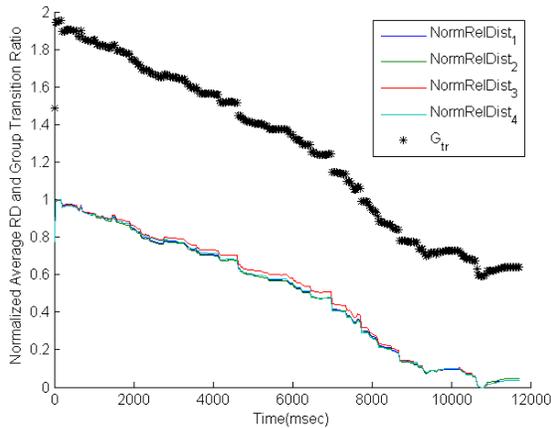
The overall workflow is shown in Figure 4, where we see the original HTD crawler and extractor with Text Analysis (unsupervised analysis using MALLET and supervised analysis using SVM in a bag-of-words model) and Image Analysis (via N-cuts followed by an unsupervised MALLET LDA analysis and a supervised analysis using SVM) modules, including fusing their results via the Fusion module and adding components to help visualize the results. A majority of the workflow fragments utilized here, including the Fusion model and the visualization modules, were re-used from earlier. The extension of the components saved effort estimated to be on the order of 300 man-hours of work. Summary results showed an Equal Error Rate of 0.37 and F-Measure of 0.47 for the Fusion module.

IV. CONCLUSION

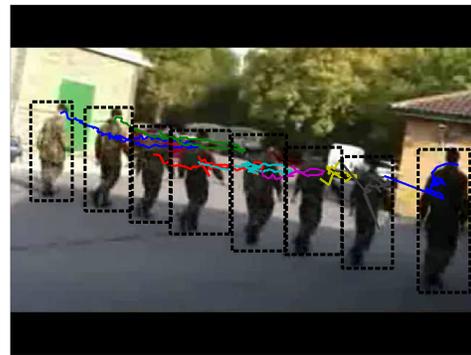
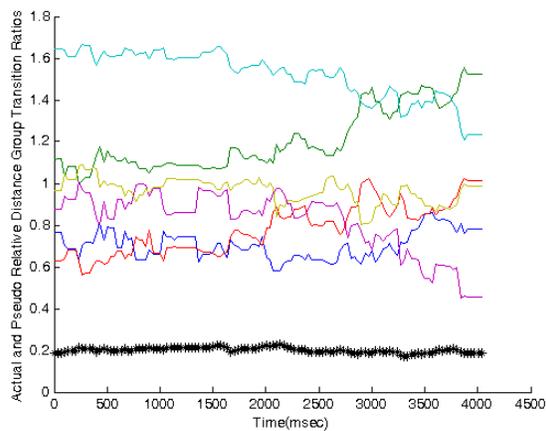
Scientific workflows thus help facilitate collaboration in computer vision by allowing for sharing, reproducibility, and extension of experiments and datasets. By providing a common, structured mechanism, WINGS provides a framework for composing algorithms and methods for computer vision analysis that is easily extensible for incorporating new algorithms and methods as well as processing additional types of data.

REFERENCES

- [1] I. J. Taylor, E. Deelman, D. B. Gannon, and M. Shields, *Workflows for e-Science: Scientific Workflows for Grids*. Springer-Verlag, 2006.
- [2] Y. Gil, V. Ratnakar, E. Deelman, G. Mehta, and J. Kim, "Wings for pegasus: Creating large-scale scientific applications using semantic representations of computational workflows," in *AAAI*, pp. 1767–1774, 2007.
- [3] D. Garijo and Y. Gil, "A new approach for publishing workflows: Abstractions, standards, and linked data," in *SC WORKS*, (Seattle, Washington), 2011.
- [4] Y. Gil, V. Ratnakar, J. Kim, P. A. Gonzalez-Calero, P. Groth, J. Moody, and E. Deelman, "Wings: Intelligent workflow-based design of computational experiments," *IEEE Intelligent Systems*, vol. 26, no. 1, 2011.
- [5] Y. Gil, P. Szekely, S. Villamizar, T. Harmon, V. Ratnakar, S. Gupta, M. Muslea, F. Silva, and C. Knoblock, "Mind your metadata: Exploiting semantics for configuration, adaptation, and provenance in scientific workflows," in *Proceedings of the Tenth International Semantic Web Conference*, 2011.



(a) *Group Formation*: G_{tr} decreases from about 2 to about 0.5. Sample image shows individuals converging and forming a group.



(b) *Group Movement*: G_{tr} stays steady at about 0.2. Sample image shows group moving together.

Figure 5: Workflow Generated Results.

- [6] Y. Gil, P. Groth, V. Ratnakar, and C. Fritz, "Expressive reusable workflow templates," in *Proceedings of the Fifth IEEE International Conference on e-Science (e-Science)*, (Oxford, UK), 2009.
- [7] J. Kim, Y. Gil, and M. Spraragen, "Principles for interactive acquisition and validation of workflows," *Journal of Experimental and Theoretical Artificial Intelligence*, 2009.
- [8] M. Hauder, Y. Gil, and Y. Liu, "A framework for efficient text analytics through automatic configuration and customization of scientific workflows," in *In Proceedings of the Seventh IEEE International Conference on e-Science*, (Stockholm, Sweden), 2011.
- [9] M. Hauder, Y. Gil, R. Sethi, Y. Liu, and H. Jo, "Making data analysis expertise broadly accessible through workflows," in *In Proceedings of the Sixth Workshop on Workflows in Support of Large-Scale Science (WORKS)*, held in conjunction with SC, 2011.
- [10] T. M. Kurc, S. Hastings, V. S. Kumar, S. Langella, A. Sharma, T. Pan, S. Oster, D. Ervin, J. Permar, S. Narayanan, Y. Gil, E. Deelman, M. W. Hall, and J. H. Saltz, "High performance computing and grid computing for integrative biomedical research," *Journal of High Performance Computing Applications*, vol. 23, no. 3, pp. 252–264, 2009.
- [11] R. Sethi and A. Roy-Chowdhury, "Physics-based activity modelling in phase space," *ICVGIP*, 2010.
- [12] R. J. Sethi, "Towards defining groups and crowds in video using the atomic group actions dataset," in *IEEE International Conference on Image Processing (ICIP)*, 2015.
- [13] B. Song, R. J. Sethi, and A. K. Roy-Chowdhury, "Robust Wide Area Tracking in Single and Multiple Views," in *Visual Analysis of Humans* (T. B. Moeslund, L. Sigal, V. Krüger, and A. Hilton, eds.), pp. 1–18, Springer-Verlag, 2011.