NIST Smart Data Flow System II – Speaker Localization

Antoine Fillinger Lukas Diduch Imad Hamchi Stephane Degre Vincent Stanford
Fantoine@nist.gov Ldiduch@nist.gov Ihamchi@nist.gov Sdegre@nist.gov Stanford@nist.gov

National Institute of Standards and Technology (NIST)
100 Bureau Dr – MS 8940
Gaithersburg, MD 20899-8940
USA

ABSTRACT
Multimodal applications require the acquisition and processing of massive amounts of information from multiple sensors. Because this process is beyond the capabilities of a single machine, we developed a sensor network data transport system that can employ the numerous computing devices required to perform the task. The computational and data acquisition processes are thus allocated to components spread across a network of systems. We present the NIST Smart Data Flow System II, which is a middleware layer that represents applications as data flow graphs, and transports information streams between the different computational components of the application.

Categories and Subject Descriptors
C.2.4 [**Computers-Communication Networks**]: Distributed Systems – *Distributed Applications.*

General Terms
Algorithms, Performance, Design, Standardization.

Keywords
Data Streams, Distributed Computing, Smart Data Flow, Smart Spaces, Multimodal Data Transport Infrastructure.

1. THE NIST SMART DATA FLOW SYSTEM II
The NIST Smart Data Flow System II (NSDFS II) is a data transport abstraction which replaces the version I, but enhances its scalability, cross-platform portability, fault-tolerance, throughput rates, and general robustness under load. NSDFS I has been successfully deployed in the NIST meeting room [1] and collects more than 200 gigabytes of data per hour from up to 287 sensors. The NSDFS II can be used to support real-time multimodal applications. This data flow system is dynamic, so computing nodes can join or leave the Smart Flow network at any time.

In addition, the NSDFS II is portable across Linux, Mac OS X, and Microsoft Windows. Data can be transported independently between these three operating systems.

The system was developed in C++ which has an object-oriented architecture. The language binding capabilities allow clients nodes to be developed using C++, or in Java using our wrapper. The NSDFS II API is used in the program in order to exchange data on a Smart Flow network.

In order to control the data flow, some convenient tools are provided by the system. The Control Center is a GUI, which allows the creation and control of multimodal applications. It allows the user to connect client nodes through flows and associate each client node to a computing device. Client nodes use the same API to exchange data within the same computation node or between nodes over the network. To give an extra level of convenience, the data flow system can be controlled from any computer using a Java applet provided for the purpose. The computer connecting to the applet doesn't need an NSDFS II installation in order to manage an NSDFS II application. An Internet browser and a current release of Java are the only things needed to use this applet.

2. THE MULTIMODAL APPLICATION
We demonstrate a speaker localization application written using the NSDFS II. This demonstration runs on three laptop computers using two sensors: a microphone array [2] and a web camera. The sensor information streams are processed to localize the person who is speaking (see figure 1).

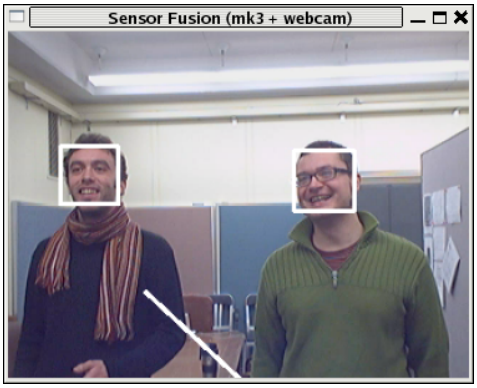

Figure 1. Speaker localization display

This multimodal application acquires data from the web camera and performs face localization on the persons, who are in view. A

IPSN'07, April 25-27, 2007, Cambridge, Massachusetts, USA.
ACM 978-1-59593-638-7/07/0004.

beamforming algorithm is simultaneously applied to data captured from the microphone array in order to estimate the bearing of the speaker. This is done using a steered-response beam former and selecting the bearing of the beam with the highest average energy. The direction of the sound and the coordinates of the faces detected in the video are correlated in order to determine accurately which person is speaking, and the speaker's face is marked by an arrow.

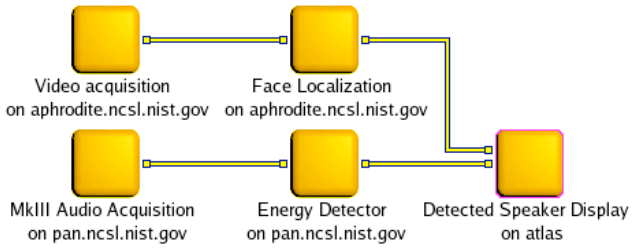

Figure 2. The speaker localization application

This localization is performed with independent programs (client nodes), which use the real-time capabilities of the data flow system in order to transport the data stream, either locally on the same computer or remotely via the network (see figure 2). Some of these client nodes acquire the data from the sensors and make them available on the Smart Flow network as a flow. Others programs consume the data transported by the NSDFS II and apply beamforming or face localization algorithms and provide their classification results encapsulated in flows. Some can just consume these flows, correlate the information, and display a result. Client nodes can join or leave a Smart Flow application at any time, providing or consuming flows, which can contain any kind of data and metadata. These client nodes can then be reused to build more complex multimodal application flow graphs.

3. DISCLAIMER AND STATEMENTS

The NIST Smart Data Flow System II software was developed at the National Institute of Standards and Technology (NIST) by employees of the Federal Government in the course of their official duties. Pursuant to title 17 Section 105 of the United States Code this software is not subject to copyright protection and is in the public domain.

Certain commercial products may be identified in order to adequately specify or describe the subject matter of this work. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products identified are necessarily the best available for the purpose.

The Smart Data Flow is an experimental system. NIST assumes no responsibility whatsoever for its use by other parties, and makes no guarantees, expressed or implied, about its quality, reliability, or any other characteristic.

The National Institute of Standards and Technology and the Smart Spaces Project would appreciate acknowledgements if the tools are used.

4. REFERENCES

[1] Stanford, V. Garofolo, J. Galibert, O. Michel, M. Laprun, C. *The NIST Smart Space and Meeting Room projects: signals, acquisition annotation, and metrics.* Proc. International Conference on Acoustics, Speech, and Signal Processing, 2003. Volume 4, 6-10 April 2003 Page(s):IV - 736-9 vol.4

[2] Xu, R. Mei, G. Ren, Z. Kwan, C. Aube, J. Rochet, C. Stanford, V. *Speaker Identification and Speech Recognition Using Phased Arrays.* Ambient Intelligence in Everyday Life, Springer Lecture Notes on Artificial Intelligence 3864, pp 227-238, 2006.