Analysis of video images used to study gas-liquid transfer

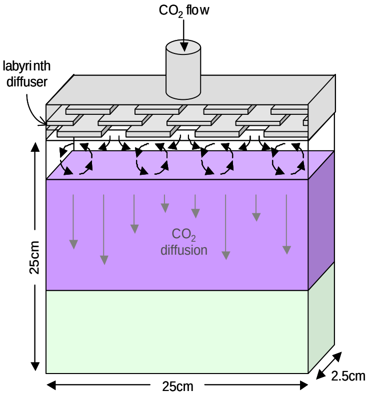
S.P.Gurden[1], E.M.Lage[2], C.G.de Faria[1], I.Joekes[2], M.M.C.Ferreira[1]

[1]Theoretical and Applied Chemometrics [2]Applied Physical Chemistry
Department of Chemistry, State University of Campinas (UNICAMP), Brazil
website: http://lqta.iqm.unicamp.br *e-mail:* spgurden@iqm.unicamp.br

Introduction

Video images and **multivariate models** are used here to provide information about the chemical changes occurring within a **CO_2-water exchange** system. This process is of interest due to its importance in terms of the **global carbon cycle**. The methodology described here could be useful for any processes in which **spectral imaging** is used to provide **dynamic information** about **spatially inhomogeneous** chemical processes.

A glass tank contains a saline solution in which a **pH indicator** is present. CO_2 gas is introduced at the top of the tank and gradually **dissolves** and **diffuses** through the solution, causing a **color change** from green (pH 6.2) to violet (pH 4.2). The entire experiment is filmed using a standard **color video camera**.

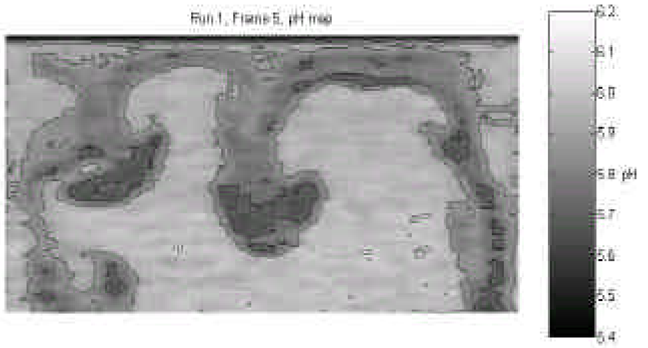


Prior to analysis using models such as PCA and PARAFAC, it is necessary to automatically **digitize**, **align** and **crop** the raw data to produce a set of **congruent** images. Each experiment produces a **multivariate movie**: a **four-way array** with dimensions **height ? width ? wavelength ? time**.

Analysis of single images

PCA & PARAFAC: Single multivariate images (height ? width ? wavelength) can be analyzed in order to understand **chemical distribution**. **PCA** treats each pixel as a separate object, while **PARAFAC** attempts to find linear structure within the image plane by modeling the *height* and *width* dimensions separately. Both approaches were found to have advantages and limitations in terms of understanding spatial distribution. However, it is notable that independent of how the spatial modes were modeled, both PCA and PARAFAC gave identical loadings in the *wavelength* mode, clearly describing the distinction between the green (low CO_2) and violet (high CO_2) regions in the color images.

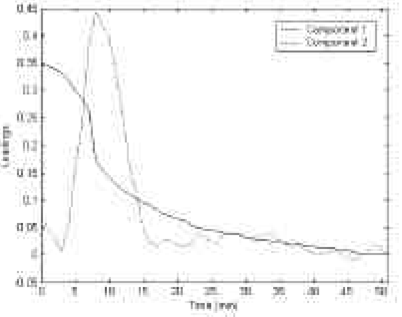
MAPPING: Another approach is to build an external **PLS calibration model** relating the spectral measurements at each pixel to a chemical property of interest, e.g. pH or [CO_2]. This reduces the dimensionality of the data by collapsing the *wavelength* mode. Here, color images at known pH's were used to build **pH maps**:



The frame above is taken from t=5min into one experimental run, The CO_2 uptake occurs from the top (gas-liquid interface) downwards, but is seen to occur inhomogeneously.

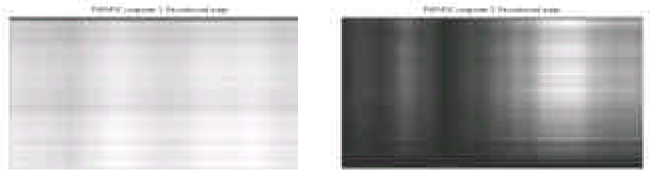
Analysis of movies

To reduce the dimensionality of the data, all multivariate movie frames were transformed into pH maps, leading to a three-way movie array (height ? width ? time). Again, this array can be analyzed using PCA (performed by first unfolding the three-way array) or PARAFAC. In this case, the **PARAFAC** model was found to be superior, giving two significant components (explaining 98.82%). The *time* mode loadings show that the first component describes an effect **general to the process**: the mean transport of CO2 from a thin layer at the liquid surface to the rest of the tank. The second component describes an **effect specific to the experiment**: a diffusion pattern in which the CO2 uptake occurred in vertical bands mostly at the left-hand-side of the tank.



Component 1 describes the flux of CO2 from the surface layer to the rest of the tank, arriving (t=52min) at a homogeneous distribution.

Component 2 describes a spatial feature due to the uneven uptake pattern of CO2 throughout the tank during t=5-15min.



After performing a PARAFAC or PCA analysis on multivariate images, it is possible to generate reconstructed loadings images, in order to gain spatial interpretation.

Analyses of other experimental runs exhibited the same pattern: one component describing the boundary layer and another describing a particular spatial feature of the CO2 uptake. This demonstrates how imaging could be used for **on-line monitoring** of chemical processes, *e.g.* detecting when an **homogeneous mixing state** has been reached or whether an **abnormal distribution pattern** is occurring. The *time* mode loadings could also be used to determine kinetic constants such as **diffusion rates**.

The **PCA** analysis gave a less interpretable model. probably because the PCA components are forced to be orthogonal. This work demonstrates that PARAFAC provides an **alternative**, and in some cases **complementary**, approach to PCA for multivariate image analysis. The fundamental difference is that PARAFAC models structure *within* the image plane, which was found to be appropriate in this case.

Conclusions

This example shows how multivariate models can successfully aid analysis of large image and movie arrays. **Spectroscopic imaging** is now becoming popular and efficient analysis of the huge arrays produced will be of great importance. In addition to PCA and PARAFAC, there exist a vast array of tools within the **image analysis sciences** (histograms, filtering, pattern recognition etc.) which may also be useful, either for chemical analysis itself or for reducing the size of the data array prior to multivariate modeling.

This research was funded by the State of São Paulo Research Foundation (FAPESP) and the Brazilian National Council for Scientific and Technological Development (CNPq).