



Good morning everyone, and thank you Marcello,

let me introduce myself, my name is Filippo Micheletti, I'm an electronic engineer working here at Applied Physics Institute "Nello Carrara", and I'm a PhD student in Information Engineering at the University of Siena.

In this presentation I will talk about hyperpectral data access and, in particular, about a platform developed here in the last years for hyperspectral data management and visualization, called CRISTINA.

Here you have my contacts, in particular if you want this presentation you can get it from that link, on my homepage.



Well, the outline of this talk is the one that you can see in this list, we will have an overview on hyperspectral data, speaking about characteristics and problems related to such kind of data, then we will speak about hyperspectral imaging softwares, not about a specific one but more generally about them to understand what problems they have related to different kind of users.

This part will be necessary to introduce our platform, which, as the slide says, represents a new approach to hyperspectral data.

So we will see the characteristics of CRISTINA both from the data-administrator and the user sides with some practical examples.

This talk will be closed by a summary of perspectives, developments and conclusions and, at the end, I will answer to a question which a lot of people asked me about this work: why the name CRISTINA? It will be a little surprise specially for Marcello...

This presentation comes first than the practical demonstration that you will have after the coffee break when Marcello will illustrate to you a practical artwork investigation using this system.



Before starting the talk I want to say thatyou can have an access to our system using a temporary account mado on the purpose of the Training School.

Credentials are the one you can read on the slide and the account will expire on Friday, at the end of the Training School

BUT

If you like our work and you think that it could be useful for your purposes please don't exitate to ask for a permanent account, we will be glad to give to you!





So let's start speaking about hyperspectral data.

Storically hyperspectral surveys was born in the field of earth observations: plane and/or satellite hyper-spectral imaging techniques have been applied to the study of resources in the country as well as to soil survey and mapping.

More recently these techniques also have been used in other applications fields, like medicine, industrial material study and, last but not least, in diagnosis and documentation of works of art mainly focused on paintings.



So let's start speaking about hyperspectral data.

Storically hyperspectral surveys was born in the field of earth observations: plane and/or satellite hyper-spectral imaging techniques have been applied to the study of resources in the country as well as to soil survey and mapping.

More recently these techniques also have been used in other applications fields, like medicine, industrial material study and, last but not least, in diagnosis and documentation of works of art mainly focused on paintings.



So let's start speaking about hyperspectral data.

Storically hyperspectral surveys was born in the field of earth observations: plane and/or satellite hyper-spectral imaging techniques have been applied to the study of resources in the country as well as to soil survey and mapping.

More recently these techniques also have been used in other applications fields, like medicine, industrial material study and, last but not least, in diagnosis and documentation of works of art mainly focused on paintings.



Hyperspectral surveys are infact the most important activity of our lab, where several kind of investigations are performed.

Here you can see for example Marcello doing Fiber Optic Reflectance Spectroscopy analysis, and two pictures of the scanner assembled by Andrea Casini and Lorenzo Stefani, applied on the right picture on the painting of Madonna dei Fusy by Leonardo da Vinci.



Hyperspectral measurements involve dealing with several issues and most of these are related to the data size.

Hyperspectral surveys in fact usually produce huge quantities of information which are arranged in files with very big dimensions.

Such big dimensions lead to problems with data distribution, storage and access.

Also the complexity of data often gives problems with the computations needed to use the data.

Another important characteristic, related mostly to the use of the data than to the data itself, is the dispersion of them. As we will see dispersion is often due to the property of the subject investigated and this leads to problems with data distribution, access and comparison.



More in detail we can say that the biggest problem is for sure data distribution and access.

Obtained data are typically arranged in cube-files, structures which contain the spectra of each point of image. Depending on both spatial and spectral resolutions of the instrumentation used and on the dimension of the scanned surface, the resulting file can be very large.

For example, the hyper-spectral imaging system assembled at IFAC-CNR is able to acquire cube-images at very high spatial (about 270 dpi on the object plane) and spectral (over 400 bands) resolution in the 400-900 nm range, which produces files of about 100 GB for squared meter.

Moving such big data is usually a problem because there are not cheap media supports able to contain so much data and usually the only way is to use a dedicated hard disk.

On the other hand different kind of data transfer could be possible, like for example through the internet, but again often connections are not so strong and affordable and either way this solution would require a lot of time to transfer data from the source to the user.

Often the solutions adopted are to undersample the data or to get only some portions of them: this is obviously not the best solution because it invalidates the quality of the survey.



It's clear that storing such big quantities of data is also a problem: if only one measurement fills an entire hard disk one can easly imagine that storing the results of several surveys requires unusual structures which not everyone can have.

Often also to give a proper organization to data is difficult and can become a real challenge.

Of course this is a problem specially for labs, but at the moment there are not efficient solutions to organize in a clear end easy way this kind of data.



Last two key words on issues with hyperspectral data are Computation and comparison.

Computation by itself could be not required but, being the data an aggregation of images, usually at least a little step of computation is required also for simpler actions, like just viewing an image.

Anyway the more interesting operations usually requires some computation, more or less heavy, but they require it.

It's obvious that working with these data requires computers with advanced computation capabilities, again for the same reasons of size and complexity.

And at the end we speak about Comparison, in the meaning of comparison between different measurements taken on the same subject.

When I started working on this topic I thought that this was the most interesting thing, to be able to compare data coming from different sources and try to understand more from the whole set than from the single measurements, but I realized that this comparison is usually hard to be obtained, mainly due to the problem of spatial sincronization.

Keep in mind spatial sincronization because it will be a key word also in the next slides.



So it's clear that the problems related with hyperspectral data can be translated on hyperspectral imaging software.



And to understand why we first need to understand which kind of users hyperspectral data usually has.

How we said, from the historical footprint of hyperspectral data applications, classical users are physicians, engineers, computer scientists and, at the end, people with a background on signal analysis.

But, as we've also said, the use of hyperspectral techniques on new applciation fields, such as cultural heritage conservation, moved the interest on hyperspectral data also near to users coming from new environments, like restorers, conservatiors, architects and, in general, people who have a different background which doesn't comprise signal analysis.



On the other hand we have softwares available for hyperspectral imaging which in general are focused on data acquisition and elaboration.

These softwares are usually born in some specific field of application and they've grown with that footprint.

Also usually they are quite heavy, platform specific and often available only for some operating systems, always propertary with quite expensive licenses and, most important, they are complicated.



At the moment what we have is a lack of softwares which often lead labs to use some self-built tools, maybe neglecting standards, or to train someone to use some specific tools of one commercial software (most used is ENVI).

It would be needed instead a software which allows users to focus on their knowledge avoiding unnecessary aspects and, in this case, this means a software for who needs to have access in a simple and easy way to data, with an intuitive and clear interface, oriented mostly to data viewing, spectra analysis and measuerements comparison, and with, eventually, the options of doing some simple operations like the extraction of sub-bands, or portion of the original measurement.





With these preconditions we started developing our platform, called CRISTINA, which is the acronym of CNR Retrieval of Images from hyper-Spectral data Through Interactive Network Access.

CRISTINA is substantially a platform intended both for who produces data and who has to use them, with a particulr focus on the field of cultural assets.

The platform is so intended for storage of data and organisation, which are arguments related principally to labs, but also for presentation and access of data, principally matter of end users, and, at the end, aggregation and elaboration, two aspects which care a lot both to labs and end users because it could be a way to have a huge quantity of information available.



I will not explain in depth tecnhical aspects of my work, but let me just indicate some characteristics of the platforms.

First of all it is web-based. This is I think the most different feature from all other softwares available and has some important impications: first you will use software which is OS independent, which doesn't require installations and it's usable everywhere, from any device with an internet connection.

Data stay on some servers and these servers offer to you some services to use them. Clients which use these services stay on your device. Servers provide needed computations on data and clients provide the interface to show results.

In this way the load is divided between who uses the data and who offers them.

The system is optimized for minimizing data transfer and resource usage, both for client and servers. Resources are allocated dinamically and most computationally heavy works are saved to be used again avoiding new computations (and waits).



The storage is implemented with a network of linked databases, with a central server which tracks the position of data on other servers, manages users access and distributes the load to optimize the access to the resources.

The data can therefore be shared from different positions and the owners of data can decide what share with users.



So, about the data organization, as we said we have a central database which tracks available data.

Data are not physically located on this server, but their availability is stored here and it is organized in 3 levels:

The top level is the artwork one.

An artwork represents the classical meaning of work of art, which can be composed from one or more elements.

Elements are in fact the second level, part of the artwork of which they belong to.

At the third level we have the measurements groups, which are sets of spatially sincronized measurements performed on the same element.

The spatial sincronization plays an important role on the organization of the data because only spatially sincronised measurements can be overlapped and so compared.



In the servers where data are physically available instead data are organized by measurement types.

At the moment we can serve 3 kind of measurements:

- hyperspectral images, which are cube files containing the spectra at each scanned point
- images extrapolated from hyperspectral images
- point measurements, like FORS

Resources in local servers are linked to the central one via an unique ID which allows groups of measurements to be accessed from users directly at their location but with an unique access point, located at the address of the platform cristina.ifac.cnr.it

					CULIU
Administrator resourc	ces				
Manage system users					-
<b>Data</b> Manage data on system da	atabase				٦
Artwork type an artwork name	Madonna dei Fusi	Add artwork	Edit artwork	Delete artwork	
Element select an element	Quadro della Madonna dei Fu 💌	Add element	Edit element	Delete element	
Measurements group select a measurements group	Measured after restoration	Add group	Edit group	Delete group	
Measurement select a measurement	•	Add measurement	Edit measurement	Delete measurement	
Name					
Enter element name Thumbnail Enter thumbnail UBI					
Owner Select an owner	Museo degli Uffizi	Add Owner			
Place Select the artwork place	Piazzale degli Uffizi, 6 - Firer 🔻	Add Place			
Element type Select the element type	Pittura su tela 🔹	Add Type			
Description Enter element description					
Credits					

System administrators have an interface which allows them to manage users and add/edit or remove data registered to database.

Administrator of local server instead must only have our presentation server software and follow some guidelines to share their data.

Server softwares are scripts wrote in PHP (and so, again, platform independent) which implements all the computational load related to servers for sending data to the user client.



There are also some requirements on file formats.

In particular cube files coming from hyperspectral surveys must be in the BIP (band interleaved by pixel) format, and FORS or other point measurements must be in csv format. That's all.

The first requirement is due to the fact that this format allows a more efficient retrieve of spectra from cube file, the second is just an agreement commonly used with this kind of data.

Anyway we support who want to share its data with our system with a set of tools built specifically to data format conversion, like for example from other cube file formats to our one.

Login interface	2	COLOR & SPA CULTURAL HER	
	Welcome to CRISTINA         Car Retrieval of Images from hyper-Spectral data Through Interactive Network Access         Username         Password         Login    ask: for an account - disclaimer - © 2012 CNR-IFAC		
COST is supported by the FLI RTD Framework Programme	COSCH Training School, Florence 10-12/12/2014 COST OFFICE	ESF provides the	4

So let's see the user interface.

First of all we have a simple login interface from where one can have access to the platform.

000	<sub>■⊏</sub> Browser		COLOR & SPACE CULTURAL HEEITA	H J .GE
	Welcome to CRIS here you can display u	TINA browser all available data in our database.		
	Madonna dei Fusi	Tavole Pinocchio Mussino	Blu di Prussia	
	20ST is supported by the EU RTD Framework Programme	COSCH Training School, Florence 10-12/12/2014	COST Office Provides the Encoderation	

After login the user would be redirected automatically to a browser which allows to navigate through available artworks and elements on the platform.

I said that the user would be because at the moment there are only a few data on the platform and the browser has been replaced by only a list of them

Available artu	vorke liet	L]	
	V01K5 115t	cou	DR & SPA
Available artwor	ks and elements	CULT	URAL HE
• Beato Angelico - Madonna della	a Stella		
<ul> <li><u>Angeli</u></li> <li><u>Angeli</u></li> </ul>			
<ul> <li>Angeli di sinistra</li> </ul>			
<ul> <li>Colori di riferimento</li> </ul>			
• Dio			
<ul> <li><u>Madonna</u></li> <li>Santi</li> </ul>			
<ul> <li>Beato Angelico - Armadio degli</li> </ul>	Argenti		
<ul> <li>Anta Sinistra</li> <li>Anta Destra</li> </ul>			
<ul> <li>Leonardo Da Vinci - Madonna</li> </ul>	dei Fusi		
o <u>iviadonna</u>			
<ul> <li>Lopez</li> </ul>			
• <u>Vari</u>			
<ul> <li>Pigments tests</li> </ul>			
<ul> <li><u>Prussian blue</u></li> </ul>			
Logout			



When you click on an element of an artwork the default measurement group of that element will be automatically opened on the main viewer.

What you see in this picture is, for example, the main view of the viewer.

It is simple, and clear, to avoid unuseful waste of space on the screen and widgets which can make you loose the focus on your work.

In the background you have the default image representing the element. In this case it is an RGB picture extracted from the hyperspectral survey on the painting in the visible wavelenghts region.



Then, by default, you will find a navigation window on the right, which is the highlighted small window you can see here.

It is composed by a reduced version of the displayed default image with a small frame that locates the shown portion of it on the whole image.

On the bottom of the navigation window there is a small panel with 4 navigation buttons, two for zoom-in and zoom-out options and one to refresh the viewer.

They are bindings for the basic navigation actions you can perform on the image: dragging the big image on the background in fact allows you to move the desired portion to show, while scrolling the mouse wheel up and down will zoom in and out the image.



In the top-left corner you have then a handler to the viewer menu which is closed by default.



And finally, in the bottom left corner, you have a scale which gives to you a reference of the real size of the image at current zoom level.

It will be updated with the zoom level changes and you can drag it to where you want on the image.



But how this viewer works?

The idea of the viewer is based on an open-source project called IIPMooViewer, the reference of the project page is on the slide.

It is a client-server viewer written on C++ and JavaScript, which uses pyramidalized images to implement a fast and optimized viewer with zoom and drag capabilities.

Pyramidal images are a particular file which contains a sort of pyramid made with multiple versions of the original image at various zoom levels.

Single pictures of the pyramid are also decomposed in a grid of tiles.

What happens is that your client measures the size of the screen and tells it to the server, with the current zoom level. The server uses this information to read the right picture on the pyramid and select only the tiles matching with the area that you can visualize.

In this way only the really needed portion of the image is transfered.

When you will drag the image the client will require new tiles to the server, which will send them back, and so it rearranges the tiles it has on the screen adding only the acquired ones and removing the other ones that cannot be more visualized because out of the screen area.

The same thing happens when you zoom in or out the image: the client sends to the server the new zoom level and gives back new tiles for current area at current zoom level.

It looks simple, and, basically, it is, but managing this machinery is quite complicated.



The performances of this viewer are quite impressive and if you consider the dimensions of the data you're dealing with you can understand them.

For example loading the whole client and the main picture of the Madonna dei Fusi takes about some hundreds of milliseconds with a wide band connection or about one second with a mobile one, while you're accessing an information contained on a file bigger than 20 GB.

One side of the Armadio degli Argenti instead need about one second with a wideband connection or 2 seconds with a mobile one, to access an huge file of about 90 GB of size.



If you click on the menu handler a column on the left of the screen will open.

There you have some menu with which you can specify some settings and perform some actions.

When you have a single image for example you can set the contrast of the image by using the highlighted selector.



Contrast of image will be trimmed depending on your needs, like here for example where we set the contrast to be equal to 3.

You can set it in the range from 0 to 10.



Under the contrast slider you have 2 buttons that you can use to toggle the visibility of navigation window and scale.

Here for example we have hidden the navigation window, to save space on the man interface.



But one of the most interesting action is the blending tool which allows you to compare directly different images coming from the same hyperspectral measurement.

Let see an example of blending an image starting from the zoom on a detail of the painting Madonna della Stella from Beato Angelico.

Il you open the Images menu you can see a selector. Clicking on it you will get a list of available images coming from the hyperspectral survey which can be overlapped with the one you're currently seeing.

Click on it and select one image, for example the one taken at the wavelength of 880 nm, in the near infra-red.

Select it and click on the plus button.



The selected image will open and it will completely overlap the curren one.

But what really happens is that you have just added an image to a sort of sandwitch of displayed ones.

In fact, if you look at the menu, you will see that a controller has been added under the one already present for the base image. Playing with this controller you can set some properties of the images on the sandwitch and, in particular, you can use the sliders to blend between them.

Take in mind that the last image you open will go on the top of the sandwitch and so at the beginning it will completely hide the other ones. It's controller is instead at the end of the controllers list.



Moving the sliders knobs you can so change the opacity of each images on the stack and get the blending effect.

Here for example we have set the opacity of the near infrared image to 75 percent.



Here to 25 percent.



And here to 0 percent, which corresponds to an invisible image.



Of course you can blend among how many images you want from the available ones list.



Let see now another interesting tool: the spectra viewer.

Close the Images menu (all your images will be kept there with your settings) and open the spectra menu.

There is an icon representing a wave: this is the button wich toggle the spectra viewer, click on it and move your mouse on the image.

You will see that the pointer is now changed: you can still use the navigation actions on the image to move the zone of interest and zoom it but if you click on a point you will get the spectra on that point.



And in fact, if you click on a point, you get the spectra in a window of the viewer.

You can take multiple spectra and each of it will have a controller on the spectra menu which allows you to focus, hide, remove and select each spectra.

These action can be performed of course also using the windows with the classical graphical interface policies, also if these windows are "virtually" created on your browser viewport.



Windows can be minimized at the bottom of the screen and a map of got point will remain on the viewed image.



Please note also that the map of points where one has got a spectra is updated with zoom level and the displayed region of the image.

If you look at the controller furthermore you will see that the color of the line of the spectra in the viewing window is automatically selected to be equal to the RGB color of the picked point on the image, but you can change it just clicking on the color flag in the controller and selecting the one you want.





Let now see more in the detail how spectra are visualized.

If you look at the window containing the chart of the spectra you will see that moving the mouse pointer on the graph the position of it pojected on the ordinates axis will be tracked and the related point highlighted with a dot on the graph line.

On the top right corner instead you will see the value of the wavelength and the one of the normalized reflectance amplitude.



Now if you highligt a slice of the chart with the mouse on the vertical direction you will get a zoom of the chart on the amplitude.





If instead you highlight a region of the chart in the horizontal direction you can get a zoom on the wavelengths.



With a double click on the chart you can anyway restore the full view.



So now let me spend some words on how displayed spectra are computed starting from hyperspectral data.

If you take a spectra on a point of the picture at a low zoom level you will notice that the obtained spectra is quite smooth.



If now you zoom in the picture and take another spectra in a near point you will see that the obtained spectra is more jagged than the previous one.



This is due to the fact that when you pick a point at a certain zoom level lower than the maximum one what you pick is in the thrut not only one point on the original hyperspectral scan but a point obtained from a subsampling of it.

So, when you pick a spectra at the maximum resolution you will get exactly the spectra of that point, but when you pick a spectra at a lower resolution you will obtain an averaged spectra obtained from a squared area centered on the point you've clicked on the image.

The more the zoom is low and the more the averaged area is big.

This is also reflected on the size of the markers of the got spectra maps, to remind the level of zoom at which you have required the spectra.



And here there is a more clear comparison between the two spectra.



As we said at the beginning of the presentation the platform is intended also for a comparison of different kind of measurements.

Our lab for example does FORS surveys, which are Fiber Optic Reflectance Spectra, an "on points" measurements because the spectra are taken on single points of interest on the subject.

With our sistem you can have an access to such kind of measurement and view them inside the main viewer.

As the slide says the requirement is still the spatial sincronization: in particular on point measurements must be referred to a point on the full resolution size of the base image displayed by the viewer.



So if you look again at the menu and open the FORS one you will see a little rainbow.



Clicking on it you will enable the visualization of available measurement for the current subject.

It will produce a map on the main image made by the points of available on point measurements.



Clicking on each point will open a window containing the spectra of that point.

Also the point that you're looking will get a marker with different color to be easily find on the map.

This feature will be soon introduced also for other kind of spectra.

The behavior of windows is the same we have seen for spectra coming from the cube file.



Of course you can mix spectra coming from different kind of measurements to have a comparison, like for example on this subject which is the famous painting Madonna dei Fusi by Leonardo Da Vinci we have both hyperspectral scanned measurement and FORS one.



This is the map of FORS measurements.

Viewer feature	es: on point measuremen	
RGB PFC 16 bit Contrast	11 → 1 → 1 → 1 → 1 → 1 → 1 → 1 → 1 → 1	-0X
Spectra x: 2566, y:944 x: 959, y:808 x: 2506, y:1600	2 0.05 400 450 500 550 660 650 Wavelength (m FOR5 steethum (x= 1620, y=2200)	700 750 800 850 m]
x: 248, y:2520       x: 1576, y:2672       FORS       x: 2220, y:410       x: 2220, y:410		00 800 900
x: 090, y:10	wavelengin prin	ng X∣Imaαe _□X

And this is an example of comparison between two spectra, one coming from FORS and one from cube file, while a lot of other stuffs have been opened and visualized.





About developments we are thinking about we have principally some elaboration tools.

For example whe would like to give to some authorized users the ability to extract portions of hiperspectral survey from data, or images and make them available to other users.

Also some tools based on the use of neural networks has been discussed and they could be useful to build automatically maps on pictures or find automatically some kind of artefacts.

Then we have some tools in developments for "normal" users, like the management of sessions, with automatic save of the state of work and sotrage of user session, then a differentiation on user level access, to give for example to some trusted users particular privileges on data mangement and also an annotation tool useful for the user to put comments and notes on its work.

We would like also to build a community around the platform, both to share experience other than data and keep in touch with developers, administrators and other users about news, new data and features availables and so on.





In conclusion a new tool for hyperspectral data analysis is available on your hands.

It is still indevelopement, it has still a lot of bugs to be fixed and lacks to be filled but it exists and it needs your help to be tested and to grown so, please use it and give us a feedback!

