Canadian Forces Image Collection Digitization Plan

Serge Tremblay, Sgt (Retired), Canadian Forces Support Unit Photo Services, Ottawa, Ontario, Canada

Abstract

The aim of this document is to demonstrate the arguments that can be put forward to request funding for preservation of a large collection of negatives. NDIL is a section of the Canadian Forces Joint Imagery Centre (CFJIC) located in Ottawa, Canada and is the central library for Canadian Forces (CF) photographers

In 2005, it was requested that options for digitizing the negative collection at NDIL be presented. The discussion, recommendations and conclusion that follow are the result of the research that went into looking at various options.

Background

Since World War I, NDIL has accumulated about 1 million images documenting the history of the CF. The collection is of very high quality and value, consisting only of images from CF professional photographers - now known as Imagery Technicians. The World War II RCAF Press Liaison (PL) series is of particular value for CF history and heritage. This series of 250,000 images are captioned in the most professional manner and the images, stored as negatives, are of a very high and pristine quality.

After WW II, Image Techs continue to document the many CF operations and activities, and these are well represented in the library's collection. The collection continues to grow as up to 9000 new images, now mostly in digital format, are added yearly.

Through the years, NDIL's collection has been used for reports, presentations, and technical studies. The CF uses these archived images for many operational and public relation reasons. The collection is used by outside agencies and the media, such as the History Channel, TLC, CBC, and CTV, to recreate CF history or to report on ongoing operations. Certainly, the World War II RCAF PL series is the most demanded part of the collection by historians, filmmakers and writers.

NDIL is doing it's best to use modern conservation techniques to slow the degradation of the collection, but a lack of manpower and adequate facilities renders this task extremely difficult. A very close relationship with the National Capital Region Conservation Group has been established to learn and discuss new image conservation strategies. Also, participation in the Image Science and Technologies yearly symposium provides the necessary knowledge for digital image preservation.

The storage facility for the NDIL collection, at building M-23 of the Montreal Road NRC campus, has been upgraded to control temperature and relative humidity (RH). However, the negative envelopes and the acetates are accumulating acid created by air pollution, humidity and natural degradation. This condition is called the "Vinegar Syndrome" from the vinegar smell of certain artifacts. The acid accelerates the deterioration of the transparencies by an alarming rate and because of the age of the CF image collection is a significant problem. It is estimated that we lose 1% to 2% of our historical CF imagery every year. Many factors affect the rate of degradation, like the processing methods utilized at that time, rough handling and the level of acidity.

The Library and Archives of Canada (LAC) has a new, "state-of-the-art", archiving facility in Gatineau, and has expressed interest in housing the CF Image collection. This would address long-term preservation of the collection, but when the LAC archives a collection the donator loses ownership and a fee is charged when copies are requested. It is not in the CF's interest to take such a route, without maintaining a way to readily access the images. The CF should at least digitize this collection and then donate the artifacts and a digital copy to LAC for permanent conservation. Please note that LAC has the best Canadian conservators who help and advise the Canadian Forces constantly and expertly in its preservation efforts.

Over 10 % of the NDIL collection has been digitized during an earlier contract and since then more has been digitized to satisfy CFJIC customers' requests, bringing the total to around 10.5%. That earlier contract was valued at $490K after the final amendment, and made use of CF equipment, the KODAK Photo Imaging Workstation (PIW), purchased for approximately $50K. The final count of scanned images under that contract was 91,393. The discussion that follows discusses options for preserving/digitizing the remainder of the collection.

Discussion

There are three steps required in digitizing the collection and preparing it for long-term storage; (1) scanning the negatives, (2) recording the metadata and (3) packaging (with new sleeves) for storage.

Scanning the negatives. The National Defence Image Library (NDIL) contains approximately 1 million images. One hundred thousand images were scanned during a previous project. Of the remainder, we estimate that 150 thousand will not be scanned for the following reasons; absence of metadata, image duplications or too much damage. This leaves an estimated 750 thousand images that require digitization. About 75% of those are black-and-white and 25% are colour negatives.

To obtain a high quality image, scanning of the negatives in the collection has been carried out following guidelines from the U.S. National Archives and Records Administration, ref. F. For most negatives in the collection (4x5 black-and-white), this results in a 7 MB file. Colour negative scans result in digital files approximately 3 times larger than black-and-white, around 21 MB files. It is evident storage of the data generated from digitizing the collection is an issue that needs to be addressed. To this point digital files have been stored on archival quality CDs. As part of another CF project, NDIL will be receiving, in

Feb 06, a new server with tape backup large enough to handle the complete collection once digitized.

Metadata entry. All the negatives are either sleeved with envelopes describing the image, or have associated cards in the card catalogue. This information is vital for maintaining the usefulness of imagery in the collection. Databases are available for managing digital images, and a server version of Portfolio by Extensis is to be delivered with the new NDIL server mentioned above. For the database to be useful, the metadata associated with each image needs to be entered and tagged to its respective image. The nature of the envelopes does not allow for automated data entry, a typist would best accomplish the task. For information contained on File Cards, the option to use OCR is being explored. It was determined, to ensure the accuracy of data entered and keyword assignment, that data entry by personnel with CF military knowledge or experience is required.

Packaging. For long-term preservation of the collection, part of the solution is to replace the acidic envelopes with new, buffered (alkaline) envelopes. These buffered envelopes are available at a cost of approximately 25 cents each. Note that the old envelopes hold the majority of acids in our collection because of their porous construction. Replacing these envelopes would extend the life of the CF image collection by an estimated 100 years.

Collection Digitization Options

The NDIL staff, to estimate the person hours required to complete the task of scanning the negatives, entering the metadata and packaging the complete collection of 750,000 images, carried out a time study. The results of the time study are reported in Annex A. To set a reasonable time for completion, a limit of 5 years is used for the project estimate. Also included in Annex A are estimates from three Ottawa area image digitization contractors. The numbers in Annex A were taken into account when weighing the options given below. A task breakdown for the contractor and CFJIC staff is contained in Annexes B and C.

Option 1: Digitize on demand, and take further action to preserve the collection: This is the current situation; the collection is being digitized as clients request photographs of interest. However, "status quo" should not be an option. Additional steps are required to preserve the collection for the long term by slowing the degradation of the negatives. Replacing the storage envelopes of the negatives will extend the life of the collection (the envelopes for the complete collection will cost almost $200K.) This simple action of replacing the envelopes is a substantial undertaking, and would take several person-years to complete.

The vinegar syndrome described earlier can be slowed by cold storage. The Image Permanence Institute, ref. C, describes the benefits of cold storage but cautions that it only slows the process – deterioration will still occur. The easiest cold storage solution is a large frost-free freezer. The cost for a large, 15 ft by 20 ft, freezer is approximately $40,000. Implementation of this type of cold storage extends the life of the collection. Cold storage will also place limitations on access to the negatives, and would hamper the ability to physically search through the collection.

Option 2: Contract to digitize, enter metadata and re-sleeve negatives: As described in the tasks in Annex B, the contractor is responsible for the entire digitization effort. All negatives will be transferred, in lots, to the contractor facility. The estimates received from area contractors vary greatly. The contractors openly admitted little time would be invested to creating a firm estimate until a Request For Proposals is released. If we use the contract in ref. E as a guideline (approximately $6/image), the estimate from the Brechin Group is the most reasonable, and the total contract value for 750,000 images ($7.90/image) would be $5.9 million.

On top of the contractor effort there will still be the requirement for CFJIC staff to review the metadata. The CFJIC Library staff will do the cataloguing of the digital files and quality control. (See estimates in Annex A and Annex B for the task descriptions.)

Option 3: Contract to digitize and re-sleeve negatives: In this option, a contractor will digitize and re-sleeve negatives only, see Annex C. Again, considering the Brechin Group estimate, the contract value for 750,000 images will be $3.1 million.

CFJIC will be responsible for entering the metadata with new staff. A separate contract could be initiated to acquire this new staff and workstations. This option eliminates the metadata entry problems and will use the unit's military experience. A drawback of this option is finding adequate office space for this activity. Re-organization of our space will be needed to add 9 more workers (maximum estimated.) (See estimates in Annex A and Annex C for the task descriptions.) For estimate purposes, 9 workers at $40,000 each per year would cost $1.8 million for five years, bringing the total for this option to $4.9 million.

Option 4: Digitize collection, re-sleeve negatives and enter metadata at CFJIC: In this option, CFJIC would play the role of a contractor and hire technicians to perform the task at CFJIC. This contract would pay for needed employees to perform all aspects of the project. CFJIC would need to supply the required scanners and workstations. A benefit of this option is keeping all the negatives safely in the Unit. The main drawback again, is finding adequate office space for this major activity. A major re-organization of our working space will be needed. (See estimates in Annex A.) For estimate purposes, 17 workers at $40,000 each per year would cost $3.4 million for five years.

Table 1 – Summary of costs for each option presented.

	Cost[1]	Mil PY[2]	Civ PY	Time to complete
Option 1	$240K	2	N/A	Indefinite
Option 2	$5.9M	2	N/A	5 years
Option 3	$4.9M	2	9[3]	5 years
Option 4	$3.4M	2	17[3]	5 years

***Note 1**: Cost includes the estimate for Civ PY, but does not include the Mil PY. **Note 2:** Extra Mil PY required for the additional workload – resleeving for Option 1 and selection and packaging for Option 2.**Note 3**: Workstations for data entry are readily available; the required number of additional scanners will need to be accessed – A Sqn currently has three high-end scanners suitable for this task.*

Recommendation

It is recommended the minimum action taken is to follow Option 1 - lowest cost option for the short term – and take steps to ensure preservation of the collection. By taking action to preserve the collection, attention can be paid to the most critical parts of the collection for digitization. However, with the current staffing levels the digitization effort will progress slowly and probably not prevent loss of some of the collection to degradation. It will require an investment in conservation materials and a cold storage facility. Not accounting for the workload on NDIL personnel, this option will require $240K of funding.

Option 4, to digitize collection, re-sleeve negatives and enter metadata at CFJIC, should be considered the best choice – it sets a target of five years to complete the digitization, maintains control of the collection within the unit and has the lowest cost of the digitization options. Following the digitization effort the collection could be transferred to the Library and Archives of Canada, who have previously indicated they would assist in funding for the buffered envelopes if they receive the collection.

Conclusion

The National Defense Imagery Library is the foremost imagery collection agency in the Canadian Forces. All still imagery from CF Image Techs is collected, selected and archived in this entity for access by all the CF and the Canadian public. The CF needs to maintain ready access to this imagery. It is a valuable asset for public relations and media exposure. Retention of the collection at NDIL is the best way to maintain this ready access until a complete digital collection is created.

It is also important that the transparencies currently held in the collection are preserved, as they are vital records of CF history. The transparencies in the collection are degrading at an alarming rate. Pending any decision as to the eventual disposition of the collection, the Canadian Forces has to take action immediately to protect this important asset. NDIL is doing the impossible with its present resources to preserve this collection but more is needed immediately to complete the process.

Thank you to Captain Marc Comeau, for providing guidance in the preparation of this paper.

Acknowledgements

The author would also like to acknowledge the support and encouragements of LCol (Retired) Frank Jonker. Without their help it would have not been possible to prepare this paper for the Canadian Forces.

Annexes List:

- ***Annex A:*** Digitization of CF Image Library and Archives Time Study
  Annex B: CF Image Library and Archives Scanning Contract Project: *Contract to digitize, enter metadata and re-sleeve negatives*
  Annex C: CF Image Library and Archives Scanning Contract Project: *Contract to digitize and re-sleeve negatives*

Acknowledgements:

Thank you, to Captain Mark Comeau, for providing guidance in the preparation of this paper.

References:

A. Library and Archives Canada: Preservation Policy, December 2001
B. NARA: U.S. National Archives and Records Administration Technical Guidelines for Digitizing Archival Materials for Electronic Access, June 2004
C. Image Permanence Institute, IPI Storage Guide for Acetate Film, revised 1996

Annex A

Digitization of CF Image Library and Archives Time Study

Note: An optimal time limit for the project is estimated at 5 years for 750,000 transparencies.

Option 2: Contract to digitize, re-sleeve negatives and enter metadata:

3 well-known Ottawa area image digitization contractors provided estimates:

Brechin Group:	$7.90/image
Proulx Brothers Inc	$1.50/image
Lux Photo Inc.	$125.00/image

Option 3: Contract to digitize and re-sleeve negatives:

3 well-known Ottawa area image digitization contractors provided estimates:

Brechin Group:	$4.15/image
Proulx Brothers Inc	$1.15/image
Lux Photo Inc.	$110.00/image

*The estimates from area contractors vary greatly. The contractors openly admitted little time would be invested to creating a firm estimate until a Request For Proposals is released.

Option 4: Digitize collection, re-sleeve negatives and enter metadata at CFJIC:

A digitization test took place at CFJIC/NDIL earlier this year over a period of two weeks. Time estimates are probably on the high-end; with practice production should increase. A high quality scan was the top priority of the exercise.

Scans per hour:	12.3
Metadata entries per hours:	10.4
Estimated total person hours needed to scan 750,000 transparencies:	60,976
Estimated total person hours needed to enter metadata in 750,000 scans:	72,115
Estimated hours one person works during 1 year (220 days x 7.5hours ea):	1,650
Number of years to scan collection with one person (60,976 / 1,650):	36 years
Number of technicians needed to scan collection in 5 years:	8(7.2)
Number of years to enter metadata with one person (72,115 / 1,650):	44 years
Number of technicians needed to enter metadata in 5 years:	9(8.8)

A maximum of 17 technicians would be needed to complete the task in the suggested 5 years. Nine data entry stations would be needed. The data entry stations are readily available; A Sqn will have seven research stations attached to its new imagery server. Scanning stations are also required, ideally one per technician for a total of eight however a workflow can be devised for two technicians per scanner. A Sqn currently owns three high-end flat-bed scanners suitable for this project, and also has 35 mm slide/film scanners available.

Annex B

CF Image Library and Archives Scanning Contract Project : *Contract to digitize, enter metadata and re-sleeve negatives.*

A contractor will digitize and enter metadata of NDIL transparencies. We estimate that 70% of the collection is 4X5 black-and-white negatives. Some 5X7s may be present. The project will start with the libraries oldest negatives consisting mainly of 4X5 black-and-white transparency originals and copies. A few colour 4X5 negatives exist in this part of the collection. The project is to be completed within a 5-year period.

	CFJIC Responsibilities		Contractor Responsibilities
	NDHQ Translation Office		Prepared by Sgt Serge Tremblay 21 September 2005
		Task Responsibilities and Details	
	Task		**Details**
1	Select images for scanning	CFJIC	4X5 negatives will be scanned first on this contract. The priority is to preserve the oldest transparencies immediately.
2	Pull negatives from storage	CFJIC	Register all negatives leaving NDIL for scanning by contractor in database.
3	Package negatives in waterproof containers for transportation to contractor's site.	Contractor	Keep sequence order at all times.
4	Inspect and clean negatives	Contractor	Physically clean negatives as needed.
5	Scan Negatives	Contractor	1) Use photographic cotton white gloves in all times to handle transparencies.
		Contractor	2) Batch Scan on CREO (or other similar scanners). Adjust levels (coarse, to get in the ballpark).
		Contractor	3) Clean scanners before digitizing. Create highest quality scan as possible. Follow guidelines from NARA, ref. E.
		Contractor	4) "Save As" the CF Negative Number. Create a TIFF file and a JPEG (compression quality '8') file of each scan.
		Contractor	6) Keep old negatives envelopes in order in a separate box.
6	Re-Sleeve Negatives (Carr McLean 4X5 Clear Sealed Flap Proline Part # PL14101). Other sizes available.	Contractor	
7	Transfer CF Image Number to Acid-Free Envelope (Carr McLean Buffered envelopes 84-263 for 4X5 negatives). Other sizes available.	Contractor	Image # only on envelope as per example. Remaining Data will be captured in metadata.
8	Type Metadata - English	Contractor	DND will provide training on metadata entry on Photoshop CS IPTC. PL negatives have Caption Cards
9	Write scans on Hard Drive or other devices for transportation to NDIL.	Contractor	CFJIC IT will migrate images from Hard Drive to NDIL Server. Contractor to keep backup of the scans until directed to delete by CFJIC IT.
10	Do QC on work done		
11	Package negatives in waterproof box for transportation to NDIL site.	Contractor	Old negatives envelopes in order in a separate box. Transport scanned images to CFJIC and pick up new batch every 2 weeks.
12	Verify that all negatives are back according to database.	CFJIC	
13	Verify metadata with Cards and Negs envelopes info.	CFJIC	
			(Sauvé) Because of the enormous amount of possible keywords, Imaging Services and NDIL decided to use the keywords found in the Image caption. The problem I've seen is that the person entering the caption/keywords does not necessarily know what the correct keyword should be. For example: F-18, CF-18, CF18, CF-18A, CF-188, Hornet, McDonnell-Douglas, Boeing, etc., could all be used to describe the same airplane.
			(Tremblay) At NDIL we are used to that situation. The first set of scans (99,000) have keywords like that. We search all spellings possible. We have a Keyword list which needs completion. We estimated earlier that we could have 30,000 keywords.
14	Accuracy/Spell Check/QC Check	CFJIC	Captions and metadata are vetted for accuracy, spelling, appropriate keywords.
15	Translate Captions/Metadata to French	NDHQ	Translation will follow. If we don't have a translator in house, the captions will be translated by NDHQ.
16	Check French translations.	CFJIC	
17	Re-File Negatives in Storage.	CFJIC	
18	Re-File Caption Cards in Storage.	CFJIC	
19	Transfer Images to Gold DVD.	CFJIC	
20	Update digital images on NDIL Server.	CFJIC	
21	Enter updated Jpeg file to database.	CFJIC	
22	Enter Image in Portfolio Database.	CFJIC	
23	Image is now ready for local/Intranet use (After translations completed).	CFJIC	

Annex C

CF Image Library and Archives Scanning Contract Project Contract to digitize and re-sleeve negatives

A contractor will digitize and enter metadata of NDIL transparencies. We estimate that 70% of the collection is 4X5 black-and-white negatives. Some 5X7s may be present. The project will start with the libraries oldest negatives consisting mainly of 4X5 black-and-white transparency originals and copies. A few colour 4X5 negatives exist in this part of the collection. The project is to be completed within a 5-year period.

	CFJIC Responsibilities		Contractor Responsibilities
	NDHQ Translation Office		Prepared by Sgt Serge Tremblay 21 September 2005

Task Responsibilities and Details

	Task		Details
1	Select images for scanning	CFJIC	4X5 negatives will be scanned first on this contract. The priority is to preserve the oldest transparencies immediately.
2	Pull negatives from storage	CFJIC	Register all negatives leaving NDIL for scanning by contractor in database.
3	Package negatives in waterproof containers for transportation to contractor's site.	Contractor	Keep sequence order at all times.
4	Inspect and clean negatives	Contractor	Physically clean negatives as needed.
5	Scan Negatives	Contractor	1) Use photographic cotton white gloves in all times to handle transparencies.
		Contractor	2) Batch Scan on CREO (or other similar scanners). Adjust levels (coarse, to get in the ballpark).
		Contractor	3) Clean scanners before digitizing. Create highest quality scan as possible. Follow guidelines from NARA, ref. E.
		Contractor	4) "Save As" the CF Negative Number. Create a TIFF file and a JPEG (compression quality '8') file of each scan.
		Contractor	6) Keep old negatives envelopes in order in a separate box.
6	Re-Sleeve Negatives (Carr McLean 4X5 Clear Sealed Flap Proline Part # PL14101). Other sizes available.	Contractor	
7	Transfer CF Image Number to Acid-Free Envelope (Carr McLean Buffered envelopes 84-263 for 4X5 negatives). Other sizes available.	Contractor	Image # only on envelope as per example. Remaining Data will be captured in metadata.
8	Write scans on Hard Drive or other devices for transportation to NDIL.	Contractor	CFJIC IT will migrate images from Hard Drive to NDIL Server. Contractor to keep backup of the scans until directed to delete by CFJIC IT.
9	Package negatives in waterproof box for transportation to NDIL site.	Contractor	Old negatives envelopes in order in a separate box. Transport scanned images to CFJIC and pick up new batch every 2 weeks.
10	Verify that all negatives are back according to database.	CFJIC	
11	Pull Caption Cards from Storage and match up with negatives if available.	CFJIC	
12	Select Appropriate Keywords (English).	CFJIC	The range of possible keywords is an unresolved problem. A simple keyword list will be created and easily added to database.
			(Tremblay) At NDIL we are used to that situation. The first set of scans (99,000) have keywords like that. We search all spellings possible. We have a Keyword list which needs completion. We estimated earlier that we could have 30,000 keywords.
13	Type Metadata - English	CFJIC	Future Development - Sgt Tremblay is working with Minolta to use the new photocopiers as OCR devices. If this works we will save a lot of time re-typing caption data from file cards.
14	Accuracy/Spell Check/QC Check	CFJIC	Captions and metadata are vetted for accuracy, spelling, appropriate keywords.
15	Translate Captions/Metadata to French	CFJIC	Translation will follow. If we don't have a translator in house, the captions will be translated by NDHQ.
16	Check French translations.	CFJIC	
17	Re-File Negatives in Storage.	CFJIC	
18	Re-File Caption Cards in Storage.	CFJIC	
19	Transfer Images to Gold DVD.	CFJIC	
20	Update digital images on NDIL Server.	CFJIC	
21	Enter updated Jpeg file to database.	CFJIC	
22	Enter Image in Portfolio Database.	CFJIC	
23	Image is now ready for local/Intranet use	CFJIC	(After translations completed).