

## 第3章 調査データの修正

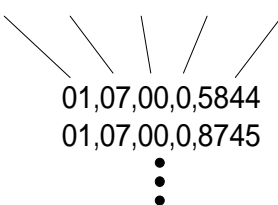
調査データは、野帳・カセットテープの記録をそのまま電子化（無修正データとする）した。しかし、この無修正データには調査記録時や電子化する際の誤りが含まれている。そこで以下のような修正が必要となる。

- 調査時の記録誤りや電子化する過程での入力誤りの修正
- マッチング処理での便宜上の修正

### 3.1 無修正データについて

無修正データは、各車線毎にファイルし、例えば片側 2 車線区間である地点番号 8 では、『8-1.TXT』（左車線）、『8-2.TXT』（右車線）というように『「地点番号」+「ハイフン」+「車線番号（1:左,2:右）」』のファイル名を付けた。また、片側 1 車線の場合には、『1-1.TXT』のように『「地点番号」+「ハイフン」+「1」』とした。

#### ● データフォーマット



01,07,00,0,5844  
01,07,00,0,8745  
⋮

...調査地点番号 (2桁)  
...車両通過時刻 (時,2桁)  
...車両通過時刻 (分,2桁)  
...車種 (1:バス,2:タクシー,0:その他一般)  
...車両のプレートNo. (大きい数字のみ,0~9999)  
(の0は読めなかった車両, ~ は各情報はカンマ区切りで保存)

時刻の入力には、正分で記入されている野帳の方を優先して使用し、時刻と野帳との関係が不明な部分はカセットテープ録音された時刻と照し合わせて補完した。

またデータは、通過した車両順に入力されている。

### 3.2 記録誤り及び入力誤りの修正

#### 3.2.1 時刻に関する誤りの修正

時刻に関する誤りには次のようなものがある。

- ある一つの調査地点の記録を通過順に見ていくと時刻の前後関係がおかしいもの。

例) 時 分 ナンバー

		:
8	18	9253
8	19	3858
		:
8	18	7727
		:

このように、1、2分単位で時刻が戻るものは野帳と照合し、野帳記録にこのような（時刻が逆になる）箇所がない場合は、電子化の際の入力誤りとして野帳記録の通りに修正し、野帳側でも時刻の逆転が確認された場合は、カセットテープ録音の記録から時刻を得て修正した。

- 一部の車両だけ時刻が不連続である。

：

10,09,55,0,5995	
10,09,55,0,9327	
10,09,05,2,0073	<b>時刻が不連続な部分</b>
10,09,55,0,0015	
10,09,55,0,6992	

このように、通過時刻が前後の車両と不連続な場合は、電子化する際の入力誤りであると考えられる。

修正の際には、野帳・カセットテープで確認後修正した。

- 数分にわたり記録のないもの。

数分間（約4分間以上）記録が途絶える部分があった場合には、必ずしも誤りであるとは言えないが記録漏れなどの可能性もあるので野帳・カセットテープで確認を行った。まず野帳にこの記録の途絶える部分がない場合には、野帳通りに再入力し、野帳にこの部分がある場合には、カセットテープで確認し必要があれば修正した。

### 3.2.2 ダブルカウントの修正

片側2車線の区間（地点8, 11, 30, 32, 34, 81, 91, 94, 95）で、図3-1のように駐車車両や車線変更などの影響で車両が車線をまたいで走行することにより、左右両方の車線の調査員が同じ車両を記録（ダブルカウント）することがあった。

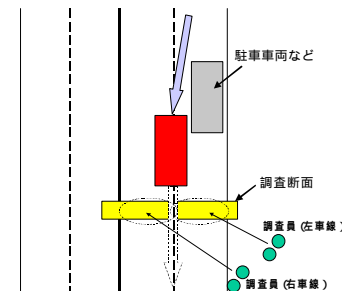


図3-1 ダブルカウント

両方の車線のデータから以下の一致条件に合ったものをダブルカウントとし片方の記録を削除した。

- **一致条件**

通過地点が一致

プレートナンバーが一致

車種が一致

2車線の通過時刻差が-1～+1分までのもの（個人誤差を考慮）

また、1車線の調査地点でも同車両を2回読む（特に混雑地点）ことが有り得るので次の条件が当てはまる複数のデータを1つのデータにした。

- 一致条件 通過地点が一致  
プレートナンバーが一致  
車種が一致  
通過時刻が一致

### 3.3 マッチング処理の利便性を考慮した修正

マッチング処理の際の利便性を考慮し、片側 2 車線ある（地点 8，11，30，32，34，81，91，94，95）地点での車線毎のデータを 1 つにまとめ地点毎のデータに変換した。この際、通過時間でソートをかけ、両車線で重複して観測されている車両を修正し 1 つにした。

### 3.4 入力時刻のズレの修正

調査に使用する時計の時刻合わせは、各監督員を中心に調査前に行うよう指示はしたものの実際には各地点毎に入力時刻のズレがあった。

まず、隣接地点でのマッチングを行い、各地点でのマッチング率（全通過車両のうち隣接した地点とのマッチングがとれた車両の割合）を調べ、マッチング率の悪い地点があった場合にはその地点または前後の地点で時刻のズレがあるものと考え個々に調べ修正した。例えば地点 80 においてはマッチングを行ったところこの地点の前後のマッチング数が極端に少なかったため地点ナンバー 80・吉祥寺駅前交差点東側東行のテープと同地点を撮影したビデオ（東京都立大撮影）をチェックしたところ、ここで調査員の時計（時計が 2 分進んでいた）が原因と考えられる入力時刻のずれが見つかった。そのため、地点 80 については、全体の時刻を 2 分早める修正を施した。

### 3.5 有効データ

修正した調査データは、調査開始直後の障害（時計が合っていなかった、調査方法を誤っていた等）があったため AM7:00～AM7:49 は使用せず AM7:50 から AM10:00 を解析に使用する有効時間とした。

この結果、総通過記録数 70999 で図 3-2 のように総通過記録数のうち車種やナンバープレートが不明だった読み取り不可は 0.4%だったため読み取り率 96.6%のほぼ全数調査と言える結果を得る事ができた。

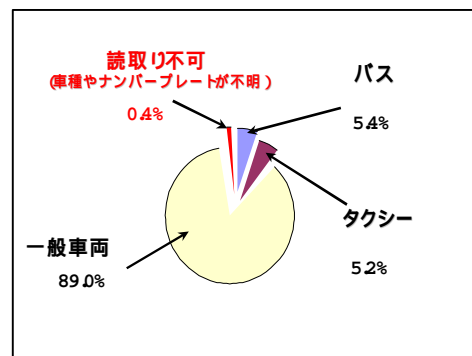


図 3-2 総通過記録数の割合